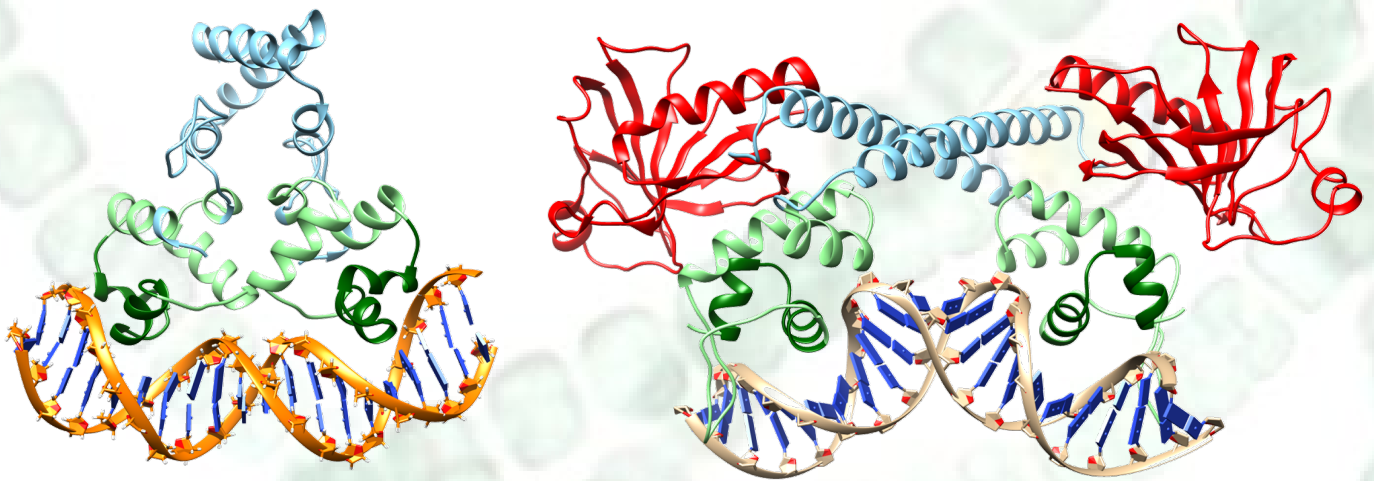


# Purificación y estudio de dos nuevos reguladores transcripcionales de la cianobacteria *Anabaena* sp. PCC 7120



Marta Acero Enguita

Trabajo Fin De Máster  
2023



Departamento de  
Bioquímica y Biología  
Molecular y Celular  
**Universidad Zaragoza**



**Universidad**  
Zaragoza



## Trabajo Fin de Máster

Purificación y estudio de dos nuevos reguladores  
transcripcionales de la cianobacteria *Anabaena* sp.  
PCC 7120

Purification and study of two new transcriptional  
regulators from the cyanobacterium *Anabaena* sp.  
PCC 7120

Autora

Marta Acero Enguita

Directores

María Francisca Fillat Castejón  
Jorge Guío Martínez

Máster en Biología Molecular y Celular  
Departamento de Bioquímica y Biología Molecular y Celular  
FACULTAD DE CIENCIAS  
2023





**Departamento de  
Bioquímica y Biología  
Molecular y Celular  
Universidad Zaragoza**

María Francisca Fillat Castejón y Jorge Guío Martínez

**CERTIFICAN**

Que la memoria titulada “Purificación y estudio de dos nuevos reguladores transcripcionales de la cianobacteria *Anabaena* sp. PCC 7120” presentada por Marta Acero Enguita ha sido realizada bajo su dirección en el Departamento de Bioquímica y Biología Molecular.

A su juicio, reúne todos los requisitos para ser presentada por su autora para la defensa de su Trabajo Fin de Máster

Zaragoza a 13 de junio de 2023

María Francisca Fillat Castejón

Jorge Guío Martínez



## Agradecimientos

Antes de abordar la memoria de este trabajo Fin de Máster, me gustaría agradecer a todas las personas que lo han hecho posible.

En primer lugar, a mis directores: a María Fillat, por la oportunidad de realizar este trabajo de fin de máster y por ser una referencia e inspiración para seguir en el mundo de la investigación y a Jorge Guío por su tutela, su orientación y su infinita paciencia en el laboratorio. También me gustaría agradecer a Emma Sevilla y Teresa Bes su ayuda y sus consejos en la realización de este trabajo, a Irene, su apoyo durante el desarrollo del mismo y a Marta Asencio, su ayuda en la realización e interpretación de los espectros de dicroísmo circular.

Finalmente me gustaría agradecer a mi familia: a mis padres y hermanas por su constante apoyo, por aguantarme durante los buenos y los malos momentos y por mantener siempre la confianza en mí.





## Índice

<b>1. Resumen/Abstract</b>	1
<b>2. Introducción</b>	3
2.1. Las cianobacterias	3
2.1.1. Características generales	3
2.1.2. Origen e impacto ecológico y ambiental	3
2.1.3. Diferenciación celular en cianobacterias	4
2.1.1. Interés biotecnológico de las cianobacterias	5
2.2. Metabolismo del nitrógeno en cianobacterias	6
2.2.1. Asimilación de nitrato y nitrito	7
2.2.2. Asimilación de nitrógeno orgánico	8
2.2.3. Fijación de nitrógeno atmosférico	8
2.2.4. Incorporación de amonio en los esqueletos carbonados	9
2.2.5. Regulación del metabolismo del nitrógeno	10
2.2.6. Mecanismos genéticos implicados en la diferenciación de heterocistos	11
2.3. Homeostasis de metales en las cianobacterias: las proteínas FUR	15
2.4. Redes de regulación transcripcional en la cianobacteria <i>Anabaena</i> sp. PCC 7120	17
2.5. Interés del trabajo	18
<b>3. Hipótesis y objetivos del estudio</b>	19
3.1. Hipótesis del trabajo	19
3.2. Objetivos	19
<b>4. Materiales y métodos</b>	20
4.1. Microorganismos y condiciones de cultivo	20
4.1.1. Cepas de <i>E. coli</i>	20
4.1.2. Condiciones de cultivo	20
4.2. Técnicas de trabajo con DNA	20
4.2.1. Extracción de DNA en <i>E. coli</i>	20
4.2.2. Reacción en cadena de la polimerasa (PCR)	21
4.2.3. Electroforesis de DNA	21
4.2.4. Purificación y cuantificación de productos de PCR	22
4.2.5. Doble digestión y ligación	22
4.3. Métodos genéticos	22
4.3.1. Preparación de células <i>E. coli</i> termocompetentes	22
4.3.2. Transformación de <i>E. coli</i> mediante choque térmico	22

4.4. Sobreexpresión de proteínas recombinantes .....	23
4.4.1. Prueba de sobreexpresión a pequeña escala.....	23
4.4.2. Prueba de solubilización.....	23
4.5. Purificación de proteínas recombinantes.....	24
4.5.1. Obtención de biomasa.....	24
4.5.2. Lisis celular.....	24
4.5.3. Purificación por cromatografía de afinidad a iones metálicos.....	25
4.6. Técnicas analíticas de proteínas.....	26
4.6.1. Precipitación con ácido tricloroacético (TCA).....	26
4.6.2. Electroforesis desnaturizante en geles de poliacrilamida.....	26
4.6.3. Cuantificación de las proteínas por espectroscopía.....	26
4.7. Caracterización bioquímica de proteínas.....	27
4.7.1. Determinación de la presencia de metales por ICP-OES.....	27
4.7.2. Ensayos de retardo en gel (EMSA).....	27
4.7.3. Dicroísmo circular.....	28
4.7.4. Análisis de oligomerización por entrecruzamiento.....	28
4.7.3. Ensayos de oligomerización mediada por puentes disulfuro.....	28
4.8. Técnicas bioinformáticas.....	29
<b>5. Resultados.....</b>	<b>30</b>
5.1. Análisis bioinformáticos de los reguladores Alr1976 y All0345.....	30
5.2. Clonaje en pET-28a.....	38
5.3. Prueba de sobreexpresión.....	40
5.4. Pruebas de solubilización.....	40
5.5. Purificación mediante cromatografía de afinidad a iones metálicos.....	41
5.6. Análisis de la estructura secundaria mediante dicroísmo circular.....	46
5.7. Determinación de la presencia de metal por ICP-OES.....	47
5.8. Análisis de oligomerización por entrecruzamiento.....	47
5.9. Análisis de oligomerización mediada por puentes disulfuro.....	48
5.10. Determinación de las condiciones de unión a DNA mediante EMSA.....	48
5.11. Estudio de la unión a regiones promotoras de genes implicados den el metabolismo del nitrógeno y la diferenciación de heterocistos.....	49
<b>6. Discusión.....</b>	<b>51</b>
6.1. Análisis bioinformático de los reguladores Alr1976 y All0345.....	51

6.2. Purificación y caracterización de los reguladores Alr1976 y All0345.....	52
6.3. Participación de Alr1976 y All0345 en la regulación de metabolismo del nitrógeno y la diferenciación de heterocistos en <i>Anabaena</i> .....	54
<b>7. Conclusiones</b> .....	56
<b>8. Bibliografía</b> .....	57



## Abreviaturas

2-OG: 2-oxoglutarato

ABC: ATP-binding cassette

Abs: Absorbancia

ATP: Adenosin Trifosfato

BLASTP: Basic Local Alignment Search Tool Protein

C: concentración

CHAPSO: 3- ([3-colamidopropilo]dimetilamonio)-2-hidroxi-1-propanosulfonato

CRP: cAMP receptor protein (Proteína receptora de cAMP)

DNA: ácido desoxirribonucleico

dNTPs: desoxirribonucleótidos trifosfato

DTT: 1,4 ditiotreitól

EDTA: Ácido etilendiaminotetraacético

EMSA: Electrophoretic Mobility Shift Assay (Ensayo de retardo en gel)

Fd: Ferredoxina

Fur: Ferric Uptake Regulator

GOE: Great Oxidation Event (Evento de Gran Oxidación)

GOGAT: Glutamato sintetasa

GS: Glutamina sintetasa

HEP: Heterocyst envelope polysaccharide (polisacáridos de la envuelta del heterocisto)

HGL: Heterocyst glycolipid (glicolípidos del heterocisto)

HTH: helix-turn-helix (hélice-giro-hélice)

ICP-OES: Inductively Coupled Plasma-Optical Emission Spectroscopy

IMAC: Immobilized metal affinity chromatography (cromatografía de afinidad a metals)

IPTG: Isopropil-  $\beta$ -D-1-tiogalactopiranosido

LB: Medio de cultivo Luria-Bertani

MW: Molecular Weight (peso molecular)

NADH: nicotinamida adenina dinucleótido

NADPH: nicotinamida adenina dinucleótido fosfato

PCC: colección de cultivos del Instituto Pasteur

PCR: reacción en cadena de la polimerasa

PDB: Protein Data Bank

PerR: Peroxide Responde Regulator

pI: punto isoeléctrico

PipX: PII interacting protein X (proteína que interacciona con PII X)

PMSF: fluoruro de fenilmetilsulfonilo

PSA: persulfato amónico

RuBisCO: Ribulosa-1,5-bisfosfato carboxilasa/oxigenasa

SDS: Dodecil sulfato de sodio

SDS-PAGE: Sodium dodecyl sulfate polyacrilamide gel electrophoresis (electroforesis en gel de poliacrilamida con dodecil sulfato de sodio)

TBE: Tris, borato, EDTA

TCA: ácido tricloroacético

TEMED: N, N, N', N' tetrametilendiamina

Tm: Melting Temperature

TPSb<sup>+</sup>: tetrafenilantimonio

Tris: Tris(hidroxiometil)aminometano

UV: ultravioleta

XRE: Xenobiotic response element (elemento de respuesta a xenobióticos)

Zur: Zinc Uptake Regulator

$\epsilon$ : Coeficiente de extinción molar

$\theta$ : elipticidad

## 1. Resumen / Abstract

### 1.1. Resumen

Las cianobacterias son microorganismos procariotas fotosintéticos con una enorme importancia ecológica, pues gracias a su capacidad fotosintética son responsables de la fijación de carbono inorgánico y uno de los principales productores primarios de oxígeno. Además, algunas especies son capaces de fijar nitrógeno atmosférico en células especializadas que reciben el nombre de heterocistos, gracias a lo cual pueden ser utilizadas como biofertilizantes o en la producción de biocombustibles. Por tanto, comprender la regulación de estos procesos es de gran interés, no solo para ampliar el conocimiento sobre la fisiología de las cianobacterias, sino también para implementar, optimizar y mejorar sus aplicaciones biotecnológicas.

Tanto la fotosíntesis como la fijación de nitrógeno tienen un alto requerimiento de metales, lo que hace el control de la homeostasis de metales sea un proceso crucial en cianobacterias, pues pese a ser esenciales su exceso genera estrés oxidativo. En procariotas este proceso está controlado por la familia de reguladores FUR (Ferric Uptake Regulator), que en la cianobacteria *Anabaena* se compone de tres parálogos: FurA, FurB y FurC. No obstante, las proteínas FUR en *Anabaena* no solo controlan la homeostasis de metales, sino que son reguladores globales que controlan muchos otros procesos, como los mecanismos de defensa al estrés oxidativo o el metabolismo del nitrógeno. Además de regular estos procesos de forma directa, se ha propuesto que pueden modular muchos otros genes de forma indirecta, orquestando una red de regulación transcripcional.

En este trabajo se han estudiado y purificado las proteínas Alr1976 y All0345, anotadas como posibles reguladores transcripcionales y cuya función es desconocida hasta el momento. Estas proteínas pertenecen a la red de regulación mediada por FurC, implicada en la modulación de la defensa a estrés oxidativo y el metabolismo nitrogenado. Mediante estudios bioinformáticos, se ha podido determinar la familia de reguladores a la que pertenecen, modelar su estructura y conocer sus relaciones filogenéticas. Además, se ha conseguido optimizar un protocolo de purificación para ambas proteínas y determinar sus condiciones de interacción con el DNA. Finalmente, dado que existían indicios que relacionaban estas proteínas con el metabolismo del nitrógeno, se estudió mediante ensayos de retardo en gel su interacción con promotores de genes implicados en este proceso.

Los resultados obtenidos muestran que los reguladores Alr1976 y All0345 se unen a regiones promotoras de genes con importantes papeles en el control del metabolismo del nitrógeno y la diferenciación de heterocistos, revelando que estos reguladores son nuevos componentes de la compleja red de regulación que controla estos procesos en cianobacterias.

## 1.2. Abstract

Cyanobacteria are prokaryotic and photosynthetic microorganisms of enormous ecological importance. Thanks to their photosynthetic capacity, they are responsible for the fixation of inorganic carbon and are one of the main primary producers of oxygen. In addition, some species are able to fix atmospheric nitrogen in specialized cells called heterocysts, and can thus be used as biofertilizers or in the production of biofuels. Therefore, understanding the regulation of these processes is of great interest, not only to expand the knowledge about the physiology of cyanobacteria, but also to implement, optimize and improve their biotechnological applications.

Both photosynthesis and nitrogen fixation have a high requirement of metals. This makes the control of metal homeostasis a crucial process in cyanobacteria, because in spite of being essential, the excess of metals generates oxidative stress. In prokaryotes, this process is controlled by the FUR (Ferric Uptake Regulator) family of transcriptional regulators, which in the cyanobacterium *Anabaena* consists of three paralogs: FurA, FurB and FurC. In *Anabaena*, FUR proteins not only control metal homeostasis, but are global regulators that control many other functional processes, such as oxidative stress or nitrogen metabolism. Interestingly, in addition to regulating these processes directly, it has been proposed that they may regulate many genes indirectly, orchestrating a transcriptional regulatory network.

In this work, the proteins Alr1976 and All0345, annotated as possible transcriptional regulators and whose function is unknown so far, have been purified and studied. These proteins belong to the FurC-mediated regulatory network, involved in the modulation of oxidative stress defense and nitrogen metabolism. By means of bioinformatics studies, it has been possible to determine the family of regulators to which they belong, to model their structure and to elucidate their phylogenetic relationships. In addition, it has been possible to optimize a purification protocol for both proteins and to determine the conditions needed for their interaction with DNA. Finally, since these proteins seemed to be related to nitrogen metabolism, their interaction with promoters of genes involved in this process was studied by electrophoretic mobility shift assays.

Results obtained in this work showed that the regulators Alr1976 and All0345 bind to promoter regions of genes with important roles in the control of nitrogen metabolism and heterocyst differentiation, revealing that these regulators are new components of the complex regulatory network that controls these processes in cyanobacteria.



## 2. Introducción

### 2.1. Las cianobacterias

#### 2.1.1. Características generales

Las cianobacterias son microorganismos procariotas fotosintéticos altamente versátiles (1). Constituyen un grupo muy ubico de microorganismos y son capaces de vivir en casi cualquier ambiente en el que la luz esté disponible, desde aguas dulces y marinas hasta ecosistemas terrestres, pasando por entornos extremos como hábitats geotérmicos, regiones polares o ambientes hipersalinos (2). Esta amplia distribución es posible gracias a su habilidad para adaptarse a diferentes tipos de estreses, como déficit de nutrientes, fluctuaciones de luz y temperatura y estrés osmótico, oxidativo o térmico (3).

El aparato fotosintético de las cianobacterias comprende tres complejos implicados en la transducción de energía y absorción de la luz: los fotosistemas I y II y los ficobilisomas (4), los cuales sirven de antena y permiten aumentar la captación de luz por parte de los fotosistemas (5). La actuación de estos complejos permite a las células fotooxidar el agua y generar poder reductor en forma de ferredoxina reducida, obteniendo así la energía necesaria para las funciones celulares (4).

Existe una amplia diversidad morfológica dentro de las cianobacterias, incluyendo especies unicelulares, especies con asociaciones coloniales o especies multicelulares que forman filamentos. Además, las células cianobacterianas pueden ser esféricas, elípticas, con forma de barril, cilíndricas, cónicas o con forma de disco (6). En función de su morfología y sus mecanismos de reproducción, las cianobacterias se clasifican en 5 secciones: I, II, III IV y V (7).

Las dos primeras secciones se componen de cianobacterias unicelulares, que pueden estar aisladas o formando agregados coloniales. La diferencia entre estas secciones, es que, en la sección I, la división se produce por fisión binaria, como sucede en *Gloeobacter*, *Gloeocapsa*, *Synechococcus* o *Synechocystis* y en la sección II, por fisión múltiple, como en *Pleurocapsa*, *Xenococcus* o *Myxosarcina*. Las secciones III, IV y V están formadas por cianobacterias multicelulares capaces de formar filamentos. Las cianobacterias de la sección III, como *Spirulina*, *Oscillatoria* o *Pseudanabaena* no son capaces de diferenciar heterocistos, mientras que las de las secciones IV y V en deficiencia de nitrógeno son capaces de diferenciar heterocistos, células especializadas en la fijación de nitrógeno. Su diferencia radica en que las cianobacterias de la sección IV, como *Anabaena*, *Nodularia*, *Cylindrospermum*, *Scytonema* o *Calothrix* solo se dividen en un plano, mientras que las de la V, como *Chlorogloeopsis* o *Fischerella* se dividen en más de un plano.

#### 2.1.2. Origen e impacto ecológico y ambiental

Las cianobacterias surgieron en el Arcaico tardío (hace unos 2900 millones de años) y jugaron un papel esencial en la transformación de la atmósfera de reductora a oxidante, en lo que se denomina “Evento de Gran Oxidación” (GOE, Great Oxidation Event) (8). Este fenómeno tuvo lugar hace unos 2450 millones de años, aunque la mayor parte de la diversidad existente de cianobacterias evolucionó después del GOE (8). A lo largo del eón Proterozoico, aproximadamente hace 1600 millones de años, surgieron los organismos eucarióticos fotosintéticos (9), mediante la endosimbiosis primaria de una cianobacteria en un eucariota unicelular seguida de procesos endosimbióticos de mayor orden (10). Desde finales del eón Proterozoico y hasta la actualidad, las cianobacterias son responsables del 20 % de la fotosíntesis de nuestro planeta, jugando por tanto un papel esencial como productores primarios de oxígeno (11).

Además de ser productoras de oxígeno, las cianobacterias contribuyen también a los ciclos biogeoquímicos del carbono y del nitrógeno. En concreto, algunos linajes cianobacterianos son capaces de fijar el nitrógeno atmosférico ( $N_2$ ) y transformarlo a formas de nitrógeno que pueden ser aprovechadas por otros organismos como amoníaco o amonio. El nitrógeno juega un papel muy importante en la productividad primaria, ya que es un nutriente limitante tanto en ambientes marinos como en terrestres, lo que hace que las cianobacterias fijadoras de nitrógeno, también denominadas diazotróficas, tengan una enorme importancia ecológica (11).

Las cianobacterias diazotróficas, a diferencia de las cianobacterias no diazotróficas y el fitoplancton procariota, pueden crecer en aguas donde el nitrógeno es limitado. Esto implica que, si además la presencia de otros nutrientes como hierro y fósforo es alta, su crecimiento se puede producir de forma excesiva, dando lugar a floraciones o *blooms* (12). Este fenómeno puede ocasionar desequilibrios en los ecosistemas, además de generar problemas en la calidad del agua, ya que las cianobacterias pueden sintetizar compuestos que alteren su olor y sabor, así como producir toxinas, que reciben el nombre de cianotoxinas (12).

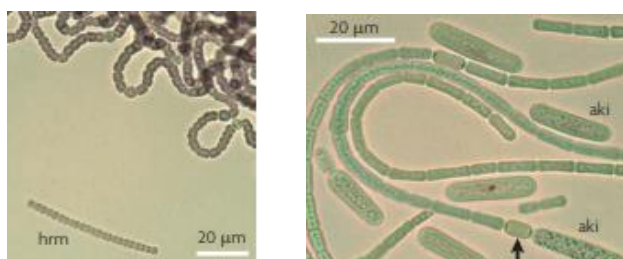
### 2.1.3. Diferenciación celular en cianobacterias

Algunas especies de cianobacterias filamentosas son capaces de diferenciar sus células vegetativas a células especializadas, entre las que se incluyen los hormogonios, los acinetos y los heterocistos, lo que constituye uno de los pocos casos de diferenciación celular en procariotas (2). En la figura 1, se muestran fotografías de las diferentes formas celulares presentes en cianobacterias filamentosas. Esta multicelularidad confiere múltiples ventajas a las cianobacterias, como son defensa frente a la depredación por protistas, mayor eficiencia en la adquisición de recursos y utilización de nutrientes, posibilidad de intercambiar información entre individuos y reducción de la interacción con competidores no cooperativos (13).

Los hormogonios son filamentos cortos que son liberados del filamento principal y que se distinguen de las células vegetativas principalmente por su motilidad deslizante y porque el tamaño de las células es menor (14). En algunas especies, los hormogonios contienen vacuolas de gas para poder regular su flotabilidad en el agua y habitualmente presentan fototaxis positiva, lo que permite la colonización de regiones iluminadas en su ecosistema (14).

Los acinetos se originan cuando el organismo está expuesto a condiciones desfavorables y son capaces de sobrevivir largos periodos de tiempo a la sequía y a altos niveles de radiación ultravioleta y germinar cuando las concentraciones vuelven a ser favorables (15).

Los heterocistos son células más grandes y redondeadas que las células vegetativas y con una menor pigmentación, así como envueltas celulares más gruesas y prominentes gránulos de cianoficina. Se forman en condiciones de deficiencia de nitrógeno combinado y permiten llevar a cabo la fijación de nitrógeno atmosférico al separar este proceso de la fotosíntesis, ya que la nitrogenasa, enzima responsable de la fijación de nitrógeno, es altamente sensible al  $O_2$  (16).



**Figura 1. Tipos celulares en cianobacterias filamentosas.** A. Imagen al microscopio de *Nostoc* sp. PCC9203, apreciándose un hormogonio (hrm) separado de las células vegetativas (17). B. Imagen al microscopio de *Anabaena* sp. PCC7120 mostrando un heterocisto marcado con una flecha negra y acinetos (aki) tanto dentro de los filamentos como libres (17).

La diferenciación de heterocistos implica un conjunto de cambios morfológicos, metabólicos y fisiológicos basados en mecanismos de expresión génica diferencial muy complejos que están encaminados a generar un ambiente anaeróbico en el que pueda tener lugar la fijación de nitrógeno. En primer lugar, desaparece el fotosistema II para evitar la generación de O<sub>2</sub>, pero no así el fotosistema I, que mediante transporte cíclico de electrones permite generar energía química en forma de ATP (15). La ausencia de fotosistema II implica que no se puede generar poder reductor en forma de NADPH, de modo que los heterocistos son dependientes de las células vegetativas para obtener equivalentes reductores. Por otro lado, se degradan las ficobiliproteínas para ser utilizadas como fuente de nitrógeno, de forma que los heterocistos tienen menor pigmentación y una apariencia más brillante (15).

Otro cambio importante en la diferenciación de heterocistos es la formación de una gruesa envuelta extracelular compuesta por dos capas químicamente diferentes situadas en el exterior de la membrana externa: una capa interna formada por glicolípidos denominada *hgl* (heterocyst glycolipid) y otra externa de polisacáridos denominada *hep* (heterocyst envelope polysaccharide) (13). La capa *hgl* sirve como una barrera de difusión de gases (15), mientras que la capa *hep* tiene un papel protector de la capa *hgl* (13). Los heterocistos son muy dependientes de las células vegetativas vecinas, pues además de necesitar equivalentes reductores requieren una fuente de carbono, ya que no expresan la enzima RuBisCO para garantizar que las reservas energéticas y de poder reductor se destinen a la fijación de nitrógeno. Por tanto, los heterocistos reciben esqueletos carbonados de las células vegetativas mientras que las células vegetativas obtienen de los heterocistos el nitrógeno fijado en forma de glutamina (16). Este intercambio de metabolitos entre células vegetativas y heterocistos es posible gracias a la presencia de canales de comunicación entre células y a la existencia de un periplasma continuo, de tal forma que las moléculas exportadas por las células pueden difundir a través del periplasma y ser obtenidas por otras a lo largo del filamento (16).

#### 2.1.4. Interés biotecnológico de las cianobacterias

Aunque como consecuencia de los florecimientos o *blooms* y de la producción de cianotoxinas pueden tener impactos negativos en ecosistemas acuáticos, en otros organismos o en la salud humana, las cianobacterias tienen un gran interés biotecnológico y pueden usarse de forma beneficiosa en muchas aplicaciones. Las cianobacterias se pueden emplear con fines farmacéuticos, puesto que sintetizan metabolitos secundarios que presentan actividades biológicas de interés, como inmunosupresores, anticancerígenos, antibacterianos, anticoagulantes, antifúngicos, antivirales y antitumorales (18), así como con aplicaciones medioambientales, destacando su empleo en biorremediación, biofertilizantes y obtención de biocombustibles.

Las cianobacterias muestran un gran potencial en biorremediación, permitiendo la eliminación de contaminantes medioambientales como pesticidas, petróleo, xenobióticos y metales pesados, así como la detoxificación de aceites industriales o el tratamiento de aguas urbanas o aguas agroindustriales. Además, se pueden emplear como agentes de bajo coste en biorremediación de aguas residuales ricas en nitrógeno y fósforo, convirtiendo estos nutrientes en biomasa que puede usarse para la producción de energía (19).

La biomasa cianobacteriana se puede emplear directamente como combustible o permitir la obtención de biocombustibles como biodiesel, biogás o biohidrógeno, el cual se ha propuesto que podría ser un buen sustituto de los combustibles fósiles (19). El hidrógeno molecular ( $H_2$ ) es producido por las cianobacterias mediante dos vías diferentes. Por un lado se puede producir mediante la enzima nitrogenasa, ya que por cada molécula de  $N_2$  atmosférico que se fija produce una molécula de  $H_2$ , o mediante la enzima hidrogenasa, que a partir de dos protones y dos electrones forma una molécula de  $H_2$  (19). No obstante, aunque la producción de biohidrógeno mediante cianobacterias es una forma ecológica y rentable de producir biocombustible, ya que el cultivo de cianobacterias tiene un coste económico muy reducido, el rendimiento en la producción de  $H_2$  es todavía muy bajo (19).

Finalmente, las cianobacterias tienen un gran potencial para agricultura sostenible. Por una parte liberan sustancias extracelulares que promueven el crecimiento de las plantas, como hormonas, vitaminas o aminoácidos, además de producir una alta variedad de compuestos biológicamente activos como antibacterianos, antifúngicos y antivirales que protegen a las plantas de ciertas enfermedades (19). Además, se ha sugerido que en suelos salinos o con un alto pH, en los que la productividad agrícola es menor, el uso de cianobacterias podría permitir su recuperación, ya que estos organismos forman un estrato grueso sobre la superficie del suelo y son capaces de conservar el carbono, nitrógeno y fósforo orgánicos, así como la humedad, de forma que se mejora la permeabilidad y aireación del suelo (19). No obstante, su principal aplicación en agricultura es su uso como biofertilizantes, gracias a la capacidad de algunas especies de fijar el nitrógeno atmosférico y liberarlo al medio en formas que puedan ser asimiladas por las plantas como amonio, polipéptidos, o aminoácidos, pudiendo llegar a contribuir con entre 20 y 30 kg de nitrógeno por hectárea, lo que las convierte en buenas candidatas para reemplazar a los fertilizantes químicos (19).

En muchas de estas aplicaciones biotecnológicas juega un papel clave la capacidad de las cianobacterias para fijar el nitrógeno atmosférico, un proceso que no pueden llevar a cabo la mayoría de los organismos, lo que hace que comprender el metabolismo del nitrógeno y la diferenciación de heterocistos en cianobacterias sean campos de estudio de gran interés

### 2.2. Metabolismo del nitrógeno en cianobacterias

Las cianobacterias contribuyen de una manera muy importante al ciclo del nitrógeno de la Tierra, ya que son capaces de incorporarlo a la biosfera a través de diferentes procesos de asimilación (20). El nitrógeno puede ser incorporado a partir de diversas fuentes como nitrato, nitrito, amonio o urea. Además, algunas cianobacterias son capaces de asimilar aminoácidos, particularmente arginina y glutamina, y las especies diazotróficas pueden llevar a cabo la fijación de dinitrógeno (20).

La fuente de nitrógeno preferida por las cianobacterias generalmente es el amonio, que es la especie de nitrógeno más reducida, pues el resto de fuentes de nitrógeno deben sufrir procesos de reducción que conllevan un gasto de energía y de equivalentes reductores para poder ser incorporados en los esqueletos carbonados (21). Por ello, la presencia de amonio juega un papel crucial en el control de los procesos de asimilación y cuando este compuesto se encuentra disponible para las cianobacterias, se activan una serie de mecanismos que reprimen las vías de asimilación del resto de formas del nitrógeno (21).

A altas concentraciones y pHs alcalinos, la mayor fracción del amonio se encuentra en forma de amoniaco, el cual puede difundir de forma pasiva a través de las membranas celulares. No obstante, en el medio ambiente por lo general el amonio se encuentra a una baja concentración, de forma que se requieren permeasas específicas (Amt) para su incorporación eficiente en la célula (22).

### 2.2.1. Asimilación de nitrato y nitrito

A pesar de que el amonio puede incorporarse del medio, esta vía es minoritaria y principalmente se obtiene de la reducción de nitratos y nitritos (22). Dada la baja concentración en la que se encuentran en el medio ambiente (20), la entrada de nitratos y nitritos en la célula se produce a través de un sistema de transporte activo mediado por un transportador tipo ABC (ATP-binding cassete) (23). En concreto, el transportador que media la entrada de nitrato y nitrito en la célula es NrtABCD, el cual está codificado por los genes *nrtA*, *nrtB*, *nrtC* y *nrtD* (24). NrtA es una lipoproteína periplásmica anclada a la membrana citoplasmática que une tanto nitrato como nitrito con una alta afinidad y NrtB es una proteína transmembrana altamente hidrofóbica, la cual se ha propuesto que forma un homodímero y representa la vía de transporte de los sustratos a través de la membrana. NrtD, así como el dominio N-terminal de NrtC son dominios de unión a nucleótidos y poseen actividad ATPasa, mientras que el dominio C-terminal de NrtC muestra homología con NrtA, aportando un sitio de unión a nitrato/nitrito intracelular y con un posible papel regulador (23). No obstante, algunas cianobacterias poseen un único gen de transporte de nitrato, denominado *nrtP*, el cual en lugar de emplear ATP como fuente de energía, emplea un gradiente de  $H^+$  o  $Na^+$  (24).

Una vez en el interior de la célula, el nitrato debe ser reducido a amonio mediante dos reacciones secuenciales, las cuales son llevadas a cabo por la nitrato reductasa (*narB*) y nitrito reductasa (*nirA*) (22). Ambas son metaloenzimas y requieren de cofactores metálicos para poder realizar su función. La nitrato reductasa contiene un cluster sulfoférico formado por 4 átomos de hierro y 4 átomos de azufre ([4Fe-4S]), así como un cofactor de molibdeno, mientras que la nitrito reductasa posee como grupos prostéticos un cluster sulfoférico [4Fe-4S] y un grupo sirohemo (23). En el proceso de reducción del nitrato, en primer lugar, la nitrato reductasa cianobacteriana cataliza la reducción de dos electrones del nitrato para dar lugar al nitrito. En esta reacción se produce un flujo de electrones desde la ferredoxina reducida obtenida de la fotosíntesis (o flavodoxina) al cluster [4Fe-4S] y después al cofactor de molibdeno, donde el nitrato es reducido a nitrito. Posteriormente, en la segunda reacción, catalizada por la nitrito reductasa, los electrones se transfieren desde la ferredoxina reducida (o flavodoxina) al cluster sulfoférico de la nitrito reductasa y de ahí al grupo sirohemo, donde el nitrito es reducido a amonio (20).

En la mayoría de las cianobacterias, incluyendo *Anabaena* sp. PCC7120, los genes del transportador ABC (*nrtABCD*), de la nitrato reductasa (*narB*) y de la nitrito reductasa (*nirA*) se encuentran constituyendo un operón denominado *nir* cuya estructura es *nirA-nrtABCD-narB* (24). Este operón se expresa en altos niveles únicamente cuando el amonio no se encuentra presente en el medio (24), pues como se ha mencionado anteriormente, en presencia de amonio, este es la fuente preferida de nitrógeno para las cianobacterias (21).



### 2.2.2. Asimilación de nitrógeno orgánico

Las cianobacterias también son capaces de asimilar nitrógeno orgánico en forma de aminoácidos o urea, la cual es una de las principales fuentes en ecosistemas marinos (25). La urea puede proceder de su excreción como producto de desecho de otros seres vivos, así como ser liberada de los sedimentos en suelos agrícolas fertilizados con urea (26). No obstante, las concentraciones de urea en los ecosistemas son bajas, por lo que para su entrada en la célula se requiere la presencia de transportadores (27).

En cianobacterias, la entrada de urea en la célula tiene lugar mediante la permeasa de tipo ABC codificada por los genes *urt*, los cuales se encuentran en un cluster en el genoma, con la estructura *urtABCDE* (27). Dicho transportador está formado por las proteínas UrtA, una lipoproteína anclada a membrana capaz de unir urea, UrtB y UrtC, proteínas integrales de membrana y UrtD y UrtE que son proteínas de unión a ATP (26). Una vez en el interior celular, las cianobacterias para metabolizar la urea, emplean la enzima ureasa, que es una metaloenzima dependiente de  $\text{Ni}^{2+}$ . La ureasa está codificada por los genes estructurales *ureA*, *ureB* y *ureC* (28), pero además requiere otros 4 genes accesorios (*ureD*, *ureE*, *ureF* y *ureG*) para la biogénesis del sitio activo de la ureasa poseedor de níquel (25). La urea se hidroliza a amoníaco y ácido carbámico, el cual se hidroliza de forma espontánea en ácido carbónico y una segunda molécula de amoníaco. Finalmente, en condiciones fisiológicas, se disocia un protón del ácido carbónico y el amoníaco se convierte en amonio (26), el cual será introducido en los esqueletos carbonados.

Las cianobacterias también pueden obtener nitrógeno a partir de aminoácidos, concretamente a partir de arginina y glutamina (28), los cuales deben ser catabolizados para poder emplear el nitrógeno que contienen (29). La glutamina se metaboliza a través del ciclo GS-GOGAT (28) y genera glutamato y amonio (21), mientras que la arginina puede ser catabolizada por dos rutas diferentes. En la primera de ellas, están involucradas la enzima arginasa, y el ciclo de la urea y se genera glutamato y amonio mientras que en la segunda participa la enzima arginina deaminasa, que produce amonio y citrulina, la cual se sigue catabolizando hasta obtener amonio (29).

### 2.2.3. Fijación de nitrógeno atmosférico

Muchas cianobacterias son diazotróficas, es decir, ante una situación de deficiencia de nitrógeno combinado, pueden fijar el nitrógeno atmosférico o dinitrógeno ( $\text{N}_2$ ). Para ello, emplean la enzima nitrogenasa, la cual es capaz de reducir el dinitrógeno a amoníaco, en una reacción energéticamente muy costosa para la célula, puesto que requiere 16 moléculas de ATP y 8 donadores de electrones (generalmente ferredoxina reducida) para fijar una molécula de  $\text{N}_2$ . Es por este motivo que la fijación de nitrógeno únicamente se produce cuando no hay otra fuente de nitrógeno disponible (15).

La enzima nitrogenasa es un complejo multienzimático compuesto por la dinitrogenasa y la dinitrogenasa reductasa, codificadas por el operón *nifHDK*. La dinitrogenasa es un heterotetramero de dos subunidades  $\alpha$  y dos  $\beta$ , codificadas por los genes *nifD* y *nifK*, respectivamente, y posee dos cofactores de hierro y molibdeno y dos cluster Fe-S (conocidos como clústeres P), mientras que la dinitrogenasa reductasa es un homodímero con un grupo sulfoférico ( $[\text{4Fe-4S}]$ ) y está codificada por el gen *nifH* (30).

Una de las principales características de la nitrogenasa es que se inactiva irreversiblemente en presencia de oxígeno. Por ello un requisito imprescindible para que las cianobacterias puedan llevar a cabo la fijación de nitrógeno es que esta enzima se encuentre en un ambiente anóxico. Dado que las cianobacterias producen oxígeno durante la fotosíntesis, estos organismos han desarrollado diferentes estrategias para poder llevar a cabo la fijación de nitrógeno (15). Cianobacterias como *Crocospaera*, *Cyanothece* o *Gloeotheca* realizan una separación temporal de la fotosíntesis y de la fijación de nitrógeno. Generalmente en presencia de luz realizan la fotosíntesis mientras que en oscuridad llevan a cabo la fijación de nitrógeno, durante la cual además tienen una elevada tasa de respiración, con el objetivo de consumir el mayor oxígeno posible y así poder conseguir un ambiente anaeróbico (15).

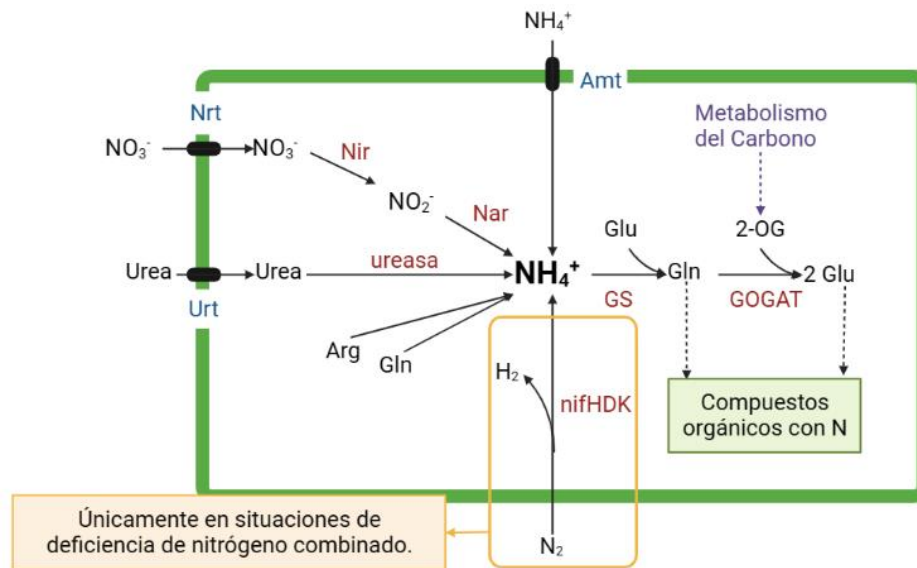
Por otro lado, cianobacterias como *Anabaena*, *Fischerella* o *Scytonema* llevan a cabo una separación espacial, mediante un proceso de diferenciación celular que da lugar a los heterocistos, células especializadas en la fijación de nitrógeno. Como se ha descrito, estas células se caracterizan por presentar una envuelta glicolípida, que impide la entrada de gases, así como por la ausencia de fotosistema II, de forma que se genera un ambiente anaeróbico (15). Finalmente, existen casos que no cumplen estrictamente ni las características de separación temporal ni las de espacial como *Trichosdemium*, en la cual la nitrogenasa se encuentra presente en 10 células consecutivas. Aunque estas células podrían considerarse diferenciadas, no presentan las características de los heterocistos, pues por ejemplo, la nitrogenasa de los heterocistos durante la noche, a pesar de funcionar a menores ratios, no es inactivada y degradada, tal y como ocurre en estas células de *Trichosdemium* (15).

#### 2.2.4. Incorporación de amonio en los esqueletos carbonados

Independientemente de la fuente de nitrógeno empleada, el amonio intracelular se incorpora a los esqueletos carbonados a través de las enzimas glutamina sintetasa (GS) y glutamato sintetasa (también llamada glutamina-oxoglutarato amidotransferasa, GOGAT), en el denominado ciclo GS-GOGAT (20). En el primer paso, la glutamina sintetasa cataliza mediante consumo de ATP la incorporación del amonio en glutamato, formando glutamina. Posteriormente, la glutamato sintetasa transfiere un grupo amida de la glutamina al 2-oxoglutarato (2-OG), una molécula intermedia del ciclo del ácido tricarboxílico (TCA), de forma que se liberan dos moléculas de glutamato (22), en una reacción que requiere poder reductor (28). A partir de las moléculas de glutamato generadas, otras aminotransferasas pueden transferir el grupo amino a otros esqueletos carbonados para generar distintos aminoácidos (22).

Respecto a las enzimas que participan en este proceso, se han encontrado dos tipos tanto de glutamina sintetasa como de glutamato sintetasa en cianobacterias. La mayoría de especies poseen únicamente una glutamina sintetasa, la cual está codificada por el gen *glnA*. No obstante, algunas especies poseen además de esta, una glutamina sintetasa hexamérica, la cual está codificada por el gen *glnN*. (21) Por otro lado, la glutamato sintetasa predominante en cianobacterias es Fd-GOGAT, la cual emplea la ferredoxina como donador de electrones y está codificada por el gen *glsF*, aunque en *Synechocystis* y *Plectonema boryanum* se ha descrito además la presencia de NADH-GOGAT, que emplea NADH como poder reductor y es codificada por *gltB* y *gltD* (22).

En la figura 2, se representan las principales vías de asimilación de nitrógeno en cianobacterias y la asimilación de amonio en esqueletos carbonados.



**Figura 2. Principales vías de asimilación de nitrógeno en cianobacterias.** Recuadrado en amarillo se encuentra la fijación de nitrógeno, proceso que se produce únicamente en ausencia de otras fuentes de nitrógeno combinado y que requiere una separación temporal o espacial debido a la sensibilidad al oxígeno de la nitrogenasa. En algunas cianobacterias filamentosas como *Anabaena*, este proceso se produce exclusivamente en los heterocistos. Adaptado de (20).

### 2.2.5. Regulación del metabolismo del nitrógeno

La activación de las vías de asimilación de fuentes alternativas de nitrógeno en ausencia de amonio se produce principalmente a nivel de expresión génica gracias al regulador transcripcional NtcA, el cual pertenece a la familia CRP (*c*AMP receptor protein) y contiene un dominio de unión al efector en el extremo N-terminal y un motivo hélice-giro-hélice de unión al DNA en el extremo C-terminal (28).

En ausencia de amonio, NtcA promueve la expresión de los genes de asimilación de fuentes alternativas de nitrógeno como los operones *nir* y *urt* implicados en la asimilación de nitrato y de urea, respectivamente, así como los genes *amt* y *glnA* implicados en la asimilación de amonio, aunque también puede actuar como represor de ciertos genes (31). El gen *ntcA* está autorregulado, mostrando un bajo nivel de expresión en condiciones de suficiencia de amonio y un aumento de su expresión, que es dependiente del propio NtcA, cuando hay una deficiencia de amonio (31). Así mismo, la transcripción de *ntcA* está modulada por el regulador de la homeostasis de hierro FurA (32) y es dependiente del regulador HetR (33).

Además, la unión de NtcA al DNA depende de los niveles de 2-oxoglutarato (2-OG) en la célula, que, en cianobacterias, sirve principalmente como esqueleto carbonado para la asimilación de nitrógeno. Ante una deficiencia de nitrógeno, como no se puede incorporar amonio a esqueletos carbonados, se produce una acumulación transitoria de 2-OG, cuyo alto nivel actúa como señal de deficiencia de nitrógeno (34) y modula la actividad de NtcA, tanto de forma directa como de forma indirecta (23). Directamente, la unión de 2-OG a NtcA potencia la unión de NtcA a sus promotores diana, ya que la unión de dos moléculas de 2-OG promueve cambios conformacionales que permiten una distancia y orientación óptima de los dominios de unión a DNA que favorecen la unión a sus promotores (22).



Por otro lado, la unión de 2-OG a la proteína PII (nitrogen regulatory protein PII) produce la liberación de la proteína PipX (PII interacting protein X) la cual queda disponible para unirse a NtcA y funciona como coactivador, posiblemente al proveer más contactos con la RNA polimerasa (35). Por tanto, aunque NtcA puede unirse por sí sola al DNA y regular la transcripción génica, los mayores cambios en la expresión de genes regulados por NtcA se producen en condiciones de deficiencia de nitrógeno ya que, como consecuencia del aumento en los niveles de 2-OG, se forma el complejo 2-OG-NtcA-PipX (22).

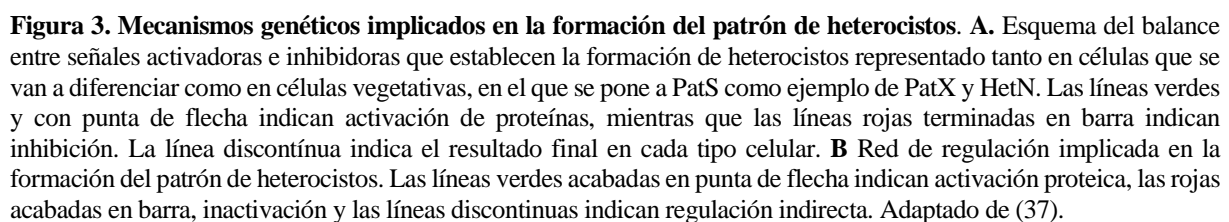
### 2.2.6. Mecanismos genéticos implicados en la diferenciación de heterocistos

Ante una situación de deficiencia de nitrógeno combinado, las cianobacterias filamentosas diazotróficas como *Anabaena*, diferencian algunas de sus células vegetativas a heterocistos, en un proceso que se puede dividir en 4 etapas: iniciación, generación del patrón, compromiso y morfogénesis. La primera etapa es la iniciación, en la que se percibe la señal de deficiencia de nitrógeno, la cual se transduce y desencadena el proceso de diferenciación. A continuación, tiene lugar la generación de un patrón, en la que se establece la distribución de los heterocistos a lo largo del filamento, de forma que queden espaciados por entre 10 y 15 células vegetativas. Posteriormente tiene lugar la etapa de compromiso, en la que diferenciación se vuelve irreversible y finalmente se produce la etapa de morfogénesis en la que se producen los cambios morfológicos y bioquímicos que dan lugar a un heterocisto funcional (36).

En la etapa de iniciación, el aumento de 2-OG es percibido por NtcA, la cual activa de manera indirecta a *hetR*, el principal regulador de la formación de heterocistos, a través del intermediario NrrA. No obstante, se han descrito otros posibles intermediarios entre NtcA y HetR, como la proteína quinasa Pkn22 y las fosfatasa PrpJ1 y PrpJ2 (37). Además, tanto *ntcA* como *hetR* están autorregulados positivamente (38). Para la correcta activación de HetR, además de NtcA, se requiere la proteína HanA, de la familia de las histonas HU, que es esencial para el desarrollo del heterocisto (39). Otra proteína necesaria para la diferenciación de los heterocistos es HetC, la cual es similar a un transportador de tipo ABC (40) y está regulada por NtcA durante las etapas tempranas de la diferenciación, aunque en estadios más tardíos también muestra regulación por HetR (41).

La distribución de los heterocistos a lo largo de los filamentos de *Anabaena* no se produce al azar, sino, que sigue un patrón, que se genera en las primeras horas en ausencia de nitrógeno. Por ejemplo, en muchas especies del orden *Nostocales*, los heterocistos se separan entre 10 y 15 células vegetativas (42). Un evento clave en la formación del patrón es la autorregulación positiva de HetR que tiene lugar en la célula que se diferencia y la inhibición de HetR por PatS en células vecinas. El gen *patS* codifica un péptido de entre 13 y 17 aminoácidos que en su extremo C-terminal posee un pentapéptido, el cual interacciona con HetR e inhibe su capacidad de unión al DNA (43). PatS no es el único inhibidor de HetR, sino que también están implicados PatX y HetN, los cuales poseen el mismo péptido que PatS para inhibir a HetR. PatX, al igual que PatS, actúa en las primeras etapas del desarrollo de los heterocistos, mientras que los niveles de HetN aumentan posteriormente, de forma que sirven de reemplazo a PatS y PatX en el mantenimiento del patrón (37). La expresión de *patS*, *patX* y *hetN* está controlada de manera indirecta por HetR y la expresión de *patX* además está controlada de forma directa por NtcA (37).

Aunque sea necesario el procesamiento de PatS para su exportación a células vecinas y para la inactivación de HetR, PatS sin procesar es activa. Por ello, para evitar la inhibición de HetR en las células que se están diferenciando a heterocistos (aquellas en las que se genera PatS) se ha desarrollado un mecanismo de inmunidad frente a PatS (42), mediado por la proteína codificada por el gen *hetL*, cuya expresión está regulada por HetR (43). HetL interacciona fuertemente con HetR sin impedir su unión al DNA y dado que interacciona con HetR en la misma región que PatS (y PatX), su unión impide la inhibición de HetR las mismas (43). Además de estas proteínas, la generación del patrón de forma correcta requiere de otros genes, a los que se les ha denominado genes *pat*, donde se incluyen *patA*, *patL*, *patN*, *patU3*, *patB*, *patC* y *patD* (44). En la figura 3, se muestra un esquema de los principales genes implicados en la formación del patrón de diferenciación.

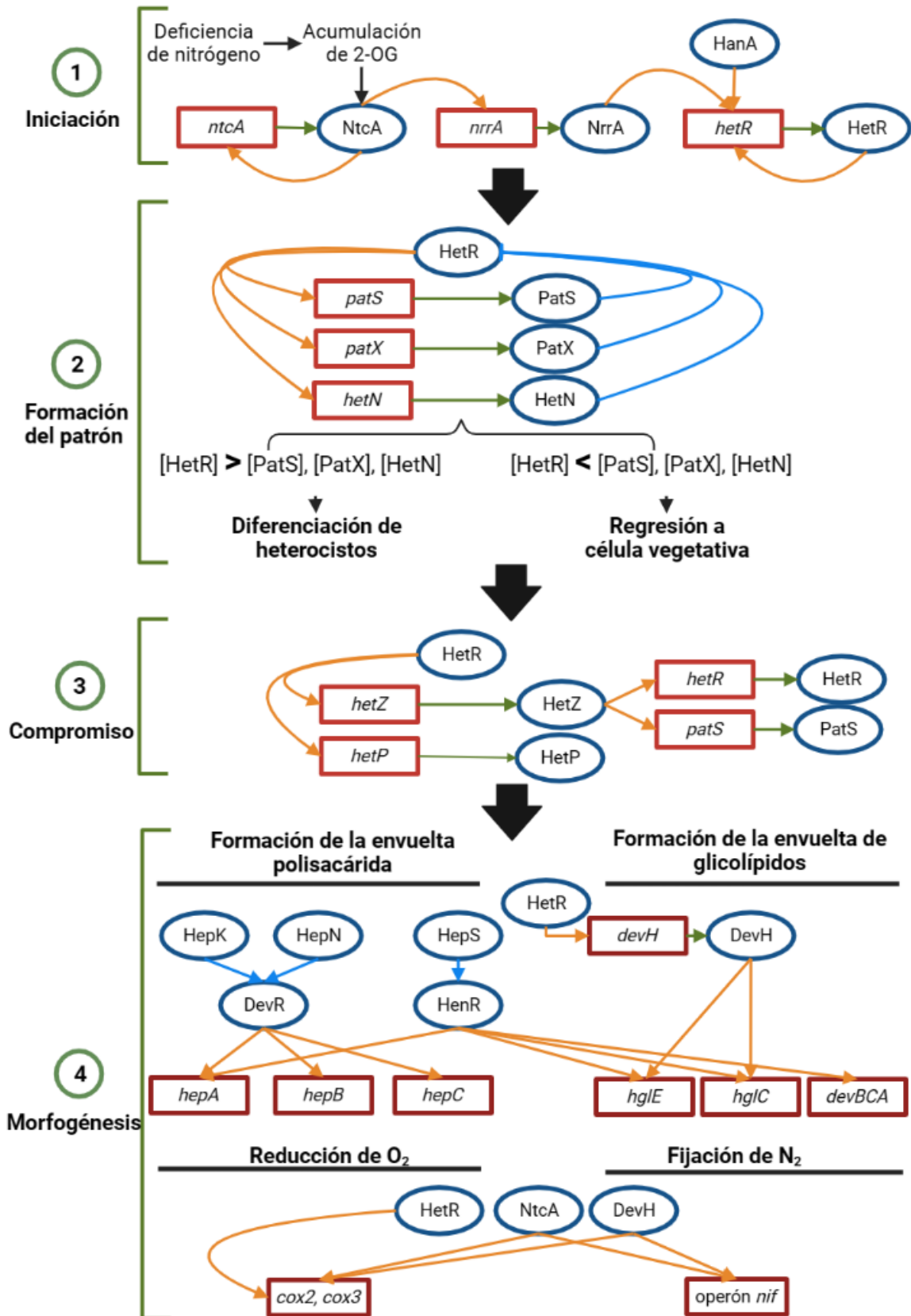


Una vez formado el patrón, se llega al punto de no retorno (etapa de compromiso), que tiene lugar tras aproximadamente 8 horas desde el inicio de la diferenciación (37) y en él, se produce la diferenciación de la célula vegetativa a heterocisto de manera irreversible (36). En esta etapa, juegan un papel clave los genes *hetZ* y *hetP*, ambos regulados directamente por HetR. HetP promueve el comienzo del proceso de diferenciación y está modulado por interacciones proteína-proteína con sus homólogos *asl1930*, *alr3234* y *alr2902*, los cuales retrasan la fase de compromiso e inhiben la diferenciación (36). Por su parte HetZ es un efector positivo de la diferenciación a través de interacciones proteína-proteína con HetR y HetP y controla la expresión de *hetR* y *patS*, amplificando la señal producida por estas proteínas, de forma que se alcancen concentraciones suficientes para llegar al punto de no retorno (45). Además se ha propuesto que HetZ y HetP tienen una función en la transducción de señal entre la fase de iniciación y morfogénesis, aunque el mecanismo molecular no es conocido todavía (37).

Una vez que se ha alcanzado el punto de no retorno, se expresan los genes requeridos para la síntesis y deposición de las envueltas celulares del heterocisto, dando lugar a la formación de las capas de glicolípidos (Hgl) y de exopolisacáridos (Hep) (37). Para la formación de la capa Hep, se requiere la expresión de *hepA*, que codifica un transportador tipo ABC (40) y está regulado directamente por HetR (46) y de las glicosiltransferasas *hepB* y *hepC* (40). La expresión de los genes *hepA*, *hepB* y *hepC* está modulada por el regulador de respuesta DevR que forma parte de un sistema de dos componentes en el que intervienen las histidín quinasa HepK y HepN y por el sistema de dos componentes de la serín treonin quinasa anclada a la membrana HepS y el regulador de respuesta HenR, el cual modula también la síntesis de la envuelta de glicolípidos Hgl (42).

Por otro lado, para la formación de la capa Hgl, se requiere la expresión de los genes *hglE*, *hglF*, *hglG*, *hglD*, *hglC*, *hglA* y *hglB*, los cuales se encuentran en un cluster y codifican enzimas asociadas con la síntesis de glicolípidos, y del gen *hglT*, que codifica una sintasa glicolípida que cataliza el paso final de la síntesis de la envuelta Hgl (37). Se han descrito dos sistemas de transporte de Hgl, uno compuesto por el transportador HgdBCD (37) y otro compuesto por un transportador de tipo ABC codificado en el operón *devBCA* (controlado por NtcA) y por HgdD, una proteína formadora de poros en la membrana externa (47). La formación de esta envuelta, además de por el sistema de dos componentes HepS/HenR, está modulada por el gen *devH*, cuya expresión depende de NtcA y HetR y que codifica un regulador transcripcional de la familia CRP que regula los genes *hglB*, *hglC*, *hglD* y *hglE* (48).

Dado que la nitrogenasa es sensible al oxígeno, durante la etapa de morfogénesis también se generan protecciones enzimáticas frente al oxígeno. En el heterocisto se expresan dos oxidasas respiratorias terminales, Cox2 y Cox3, dependientes tanto de NtcA y HetR (42) como de DevH (48), y proteínas flavin-dihierro como Flv1B y Flv3B, que reducen directamente el O<sub>2</sub> a agua sin producción de especies reactivas de oxígeno (42). En el heterocisto maduro también se expresan los genes *petH*, que codifica la ferredoxina-NADP<sup>+</sup> reductasa del heterocisto y tiene un rol importante en la distribución de electrones; *glnA*, que codifica la glutamina sintetasa, la cual incorpora el amonio al glutamato para generar glutamina; y el operón *nifHDK*, que codifica el complejo de la nitrogenasa y que está regulado por PatB (CnfR) (42). Además se producen excisiones génicas que permiten la eliminación de elementos de DNA que interrumpen los genes *fdxN* (ferredoxina del heterocisto), *nifD* (subunidad  $\alpha$  de la dinitrogenasa) y *hupL* (subunidad de la hidrogenasa) mediante las recombinasas XisA, XisF y XisC, respectivamente (49).



**Figura 4. Procesos genéticos implicados en la diferenciación de heterocistos.** Recuadrados en rojo se marcan los genes, mientras que las proteínas están rodeadas en azul. Las flechas amarillas indican regulación transcripcional, las verdes, transcripción y traducción de genes a proteínas y las azules interacción proteína-proteína. Las líneas acabadas en barra indican inhibición.

### 2.3. Homeostasis de metales en cianobacterias: las proteínas FUR

El control de homeostasis de metales es muy importante en todos los microorganismos, pero en cianobacterias tiene mayor importancia porque estos organismos tienen un requerimiento de metales mucho más alto. Esto se debe a principalmente a dos procesos celulares que tienen un alta demanda de metales: la fotosíntesis, pues los complejos fotosintéticos como los fotosistemas I y II o el citocromo  $b_6f$  son metaloproteínas que requieren gran cantidad de metal, principalmente hierro y manganeso, y el metabolismo del nitrógeno (50). Respecto al metabolismo del nitrógeno, en la asimilación de nitrato y nitrito, las enzimas nitrato y nitrito reductasa poseen grupos sulfoferricos ( $[4Fe-4S]$ ) y además la primera contiene un cofactor de molibdeno y la segunda un grupo sirohemo, que contiene hierro (23) mientras que la enzima ureasa, implicada en el metabolismo de la urea, es dependiente de níquel (25). No obstante, junto con la fotosíntesis, el proceso con un mayor requerimiento de metales es la fijación de nitrógeno atmosférico, ya que las enzimas que forman el complejo nitrogenasa requieren cofactores formados por hierro y molibdeno, y grupos prostéticos de hierro y azufre (30). Además, algunas proteínas implicadas en diferenciación de heterocistos o fijación de nitrógeno, como el regulador PatB y la ferredoxina del heterocisto FdxH, requieren hierro (32). Por ello, la homeostasis de metales, además de para otros muchos procesos celulares, es un proceso fundamental para que el metabolismo y la fijación del nitrógeno se desarrollen correctamente

Aunque la presencia de metales es crítica para el correcto funcionamiento celular, cuando su concentración en la célula aumenta tienen lugar reacciones redox indeseadas o se produce la unión de metal a sitios inapropiados de enzimas, desarrollando efectos tóxicos. Por ello, la adquisición de metales debe estar estrictamente controlada y se han desarrollado mecanismos que permiten mantener la homeostasis de metales (51). Entre los metales de transición, el hierro es uno de los factores más importantes para el correcto desarrollo de procesos fisiológicos de la célula, tales como fotosíntesis, respiración, fosforilación oxidativa, asimilación de nitrato y, sobre todo, la fijación de nitrógeno en cianobacterias diazotróficas, pues el complejo nitrogenasa requiere 40 átomos de hierro por molécula de enzima (52). El hierro, a pesar de ser el cuarto elemento más abundante de la corteza terrestre, se encuentra poco biodisponible, por lo que las cianobacterias, han desarrollado sideróforos y otros mecanismos de incorporación de hierro para poder disponer de la cantidad suficiente de este metal (53). No obstante, el hierro intracelular debe mantenerse en niveles permisivos, ya que en condiciones aeróbicas produce especies reactivas de oxígeno mediante la reacción de Fenton y resulta tóxico (53). Dado que es un elemento imprescindible para el metabolismo de las cianobacterias, pero su exceso produce toxicidad (53), su homeostasis debe estar estrictamente controlada.

En procariotas, la homeostasis de metales está controlada principalmente por los reguladores pertenecientes a la familia FUR (Ferric Uptake Regulator) (54). Por lo general, las proteínas Fur actúan como metalopresores, uniéndose en forma de dímeros a secuencias palindrómicas ricas en adeninas y timinas encontradas en los promotores de sus dianas y denominadas cajas Fur. Para ello, emplean un metal como co-represor, el cual al unirse a las proteínas FUR produce un cambio en el regulador, que promueve y estabiliza su unión con el DNA (54). Las proteínas FUR comparten un plegamiento común, compuesto por una hélice alada, que constituye el dominio de unión a DNA en el extremo N terminal y por un sitio de unión a metal que se encuentra en el extremo C-terminal, el cual además permite la dimerización de la proteína (54).



En la cianobacteria fijadora de nitrógeno *Anabaena* sp. PCC 7120, se han descrito tres parálogos de las proteínas FUR: FurA, que es el ortólogo de Fur (Ferric Uptake Regulator), FurB, que corresponde al ortólogo de Zur (Zinc Uptake Regulator) y FurC, que es el ortólogo de PerR (Peroxide Response Regulator), los cuales, además de controlar la homeostasis de metales, están estrechamente relacionados con el control del metabolismo del nitrógeno, entre otros procesos.

FurA es el principal regulador de la homeostasis de hierro, controlando los principales mecanismos de absorción de este metal por la célula (55). No obstante, FurA no solo está implicado en la homeostasis de hierro, sino que modula otros muchos genes relacionados con otros procesos fisiológicos como la fotosíntesis o la defensa frente a estrés oxidativo, y se considera un regulador global (56). Entre estos procesos, destaca su papel como regulador del metabolismo del nitrógeno y la diferenciación de heterocistos. En los heterocistos, durante las primeras 15 horas tras la deficiencia de nitrógeno, la expresión de *furA* es activada por NtcA, mientras que en células vegetativas NtcA reprime su expresión (32). Además su apropiada expresión es necesaria para el correcto desarrollo de los heterocistos, pues la desregulación de *furA* produce alteraciones en la diferenciación de heterocistos (57). Por otro lado, FurA actúa como represor de *ntcA* y se ha propuesto que podría disminuir la expresión de *ntcA* en estadios tardíos de la diferenciación de heterocistos.

Así, FurA permite integrar la homeostasis de hierro y el metabolismo del nitrógeno (57). En suficiencia de hierro, NtcA aumenta la expresión de *furA* durante la primera fase de diferenciación del heterocisto. Estos niveles aumentados de FurA en las etapas centrales y tardías de desarrollo del heterocisto regulan negativamente la expresión de *ntcA*, de manera que FurA colabora en la regulación de la inducción de NtcA durante la diferenciación de heterocistos (57). Por el contrario, en deficiencia de hierro, hay una mayor expresión de *furA*, lo que provoca una represión de *ntcA* desde el comienzo de la diferenciación del heterocisto que impide que el proceso tenga lugar (57). Además de regular *ntcA*, FurA se une al promotor de *hetR* (principal regulador de la diferenciación de heterocistos) (57) y, activa de forma directa otros genes implicados en la formación del heterocisto como *hetC* o *patA* (55).

FurB regula la homeostasis del zinc, pues controla proteínas implicadas en la incorporación de este metal en la célula, reprimiendo la transcripción de genes diana en condiciones de suficiencia de zinc (58). Además tiene un papel protector en condiciones de estrés oxidativo, pues se une de forma inespecífica al DNA protegiéndolo del daño oxidativo (59) y se ha descrito que también está implicado en otros procesos fisiológicos de la cianobacteria como fotosíntesis y respiración, metabolismo del carbono, síntesis de la pared celular, metabolismo de aminoácidos, formación de biofilms y metabolismo del nitrógeno y diferenciación de heterocistos (60). El transcriptoma de una estirpe de delección de *furB* muestra cambios en genes implicados en el metabolismo del nitrógeno, como el aumento de los transcritos de *nifK* y la disminución de la proteína de unión a ATP del transportador de nitrato, así como alteraciones en la diferenciación de heterocistos (60). Además, esta estirpe muestra menor transcripción de ciertos genes implicados en la formación del heterocisto como *trpE* (60), que codifica para la antranilato sintetasa, necesaria para inducir la diferenciación (61), así como de un homólogo de *hetP*, de *patA*, requerido para la correcta formación del patrón y de *nifD* que codifica la cadena  $\alpha$  de la nitrogenasa (60). Finalmente, se observa un aumento de la expresión de la cianoficina y la cianoficina sintetasa (60), implicadas en la degradación y síntesis de la cianoficina respectivamente, que funciona como una reserva dinámica de nitrógeno (62).

FurC se corresponde con el ortólogo de PerR y está implicado en la respuesta a estrés oxidativo, pues regula mecanismos de detoxificación en respuesta a peróxido de hidrógeno, como algunas peroxirredoxinas (52). La unión de FurC al DNA está modulada por peróxido de hidrógeno, el cual, por medio de un fenómeno de oxidación catalizada por metal, hace que se libere del DNA (63). FurC participa también en otros procesos celulares, pues está implicado en procesos morfológicos y de división celular o en la fotosíntesis (52).

Al igual que los otros parálogos FUR, FurC también participa en el metabolismo del nitrógeno y diferenciación de heterocistos (64). Su expresión está modulada por NtcA, el cual se une a su promotor (32), y además la sobreexpresión de FurC en *Anabaena*, impide que esta desarrolle heterocistos en deficiencia de nitrógeno (64). En concreto, se ha propuesto que FurC podría jugar un papel importante en el establecimiento del patrón de heterocistos pues regula genes como *asr1734* (regulador negativo de la diferenciación) y *hetZ* (efector positivo de la diferenciación) (64).

FurC también modula la biosíntesis de las envueltas del heterocisto, pues se une a los promotores del gen *hepC* y del operón *devBCA* y parece reprimir la expresión del operón *nifHDK*, el cual codifica las proteínas estructurales del complejo nitrogenasa, necesario para la fijación de nitrógeno (64). Además, el operón *xisHI*, el cual codifica para la recombinasa XisF1, es necesario para la escisión de un elemento de DNA que interrumpe el gen de la ferredoxina del heterocistos (*fdxN*), y el gen *rbrA* (ruberitrina) también están bajo el control de FurC. Finalmente, FurC regula de manera indirecta el gen *patA*, implicado en la formación del patrón, ya que se observa un aumento notable de su expresión en el mutante de sobreexpresión de FurC, pero el regulador no se une a su región promotora (64). Por tanto, además de regular directamente genes implicados en el metabolismo del nitrógeno y la formación del heterocisto, las proteínas Fur parecen regular también de forma indirecta genes que juegan papeles importantes en estos procesos.

#### 2.4. Redes de regulación transcripcional en la cianobacteria *Anabaena* sp. PCC 7120

El análisis de los transcriptomas estirpes de desregulación de las proteínas FUR en *Anabaena* sp. PCC 7120 reveló que se producía expresión diferencial de un gran número de genes con funciones reguladoras, principalmente reguladores transcripcionales y sistemas de dos componentes. Los sistemas de dos componentes son sistemas de regulación basados en un mecanismo de acoplamiento estímulo-respuesta, el cual permite a los organismos detectar y responder a cambios en diferentes condiciones ambientales. El sistema canónico consiste en una proteína histidin quinasa, la cual detecta los cambios, y activa al segundo componente, que es una proteína reguladora de respuesta, la cual suele actuar como un regulador transcripcional (65).

En concreto, en el transcriptoma de una estirpe de delección condicional de FurA se observó que se alteraba la expresión de 170 genes (56), en el de la estirpe de delección de FurB, 25 (60) y en el de la estirpe de sobreexpresión de FurC, 30 (64). De esta forma, las proteínas FUR, además de actuar como reguladores directos, llevan a cabo una regulación indirecta a través de la modulación de genes con funciones reguladoras, estableciendo redes de regulación transcripcional.

Con respecto a FurA, se ha observado que regula de forma directa tres genes que codifican reguladores transcripcionales, denominados *all1651*, *all3903* y *calB* y seis genes que codifican elementos de sistemas de dos componentes: en concreto tres reguladores de respuesta (*alr0072*, *all3348* y *rpaA*), una sensor-quinasa híbrida (*all1804*) y el operón *all7584-83*, que codifica una sensor quinasa (*all7583*) y un regulador de respuesta (*all7584*) de un sistema de dos componentes (Jorge Guío, datos no publicados). Por su parte, FurB controla la expresión dos operones que codifican sistemas de dos componentes (*all7584-83* y *all7606-05*) y genes que codifican una sensor-quinasa híbrida (*all5210*), un regulador de respuesta (*alr0072*) y una serin/treonin quinasa (*all0900*) (Jorge Guío, datos no publicados). Finalmente, FurC modula la expresión de tres genes que codifican reguladores transcripcionales: *alr1976*, *all0345* y *all1651*, así como genes de una sensor-quinasa (*all3359*) y de tres serin/treonin quinasas (*alr0354*, *alr0709* y *pknE*) (Jorge Guío, datos no publicados).

Además, se estudió la conexión de la red regulatoria de las proteínas FUR con la del regulador global del metabolismo nitrógeno, NtcA. Se encontró que una parte importante de esta red estaba corregulada por NtcA, lo que sugiere que muchos de estos genes pueden estar implicados en el metabolismo del nitrógeno (Jorge Guío, datos no publicados). En algunos casos se conoce la función de los reguladores que forman parte de esta red, como CalB que es un regulador de la familia AbrB y que junto con CalA está involucrado en la adaptación a deficiencia de nitrógeno (66). Sin embargo, para la mayoría de los elementos que componen esta red de regulación transcripcional, tanto las funciones como los procesos celulares que controlan son desconocidos.

### 2.5. Interés del trabajo

El metabolismo del nitrógeno es uno de los procesos fisiológicos más importantes en cianobacterias y está relacionado con muchas de sus aplicaciones biotecnológicas. Por lo tanto, el conocimiento detallado de la regulación del metabolismo nitrogenado en cianobacterias no solo permite comprender mejor su fisiología sino también realizar modificaciones genéticas encaminadas a potenciar o mejorar su potencial biotecnológico. No obstante, el control y regulación del metabolismo del nitrógeno, es un campo que todavía presenta muchas incógnitas.

En este trabajo se seleccionaron para su estudio dos reguladores que forman parte de la red de regulación de FurC: Alr1976 y All0345, cuyas funciones celulares son desconocidas. No obstante, se observó que el gen *alr1976* presentaba expresión diferencial en un análisis transcriptómico de *Anabaena* en condiciones de deficiencia de nitrógeno combinado frente a condiciones en presencia de nitrato tras 12 horas desde el inicio de la deficiencia (67). Por su parte, también se observó que el gen *all0345* mostraba expresión diferencial en el transcriptoma de la estirpe sobreexpresión de FurC frente a la silvestre en condiciones de deficiencia de nitrógeno combinado, pero no en condiciones de presencia de nitrato (64).

Estas evidencias sugieren que los reguladores Alr1976 y All0345 podrían estar implicados en la regulación del metabolismo del nitrógeno y la diferenciación de heterocistos en *Anabaena* sp. PCC 7120. Por tanto, su purificación y estudio son de gran interés, ya que podrían aportar información y resolver algunas incógnitas sobre el control del metabolismo del nitrógeno en cianobacterias.



### 3. Hipótesis y objetivos del estudio

#### 3.1. Hipótesis del trabajo

La hipótesis que ha dado lugar a este trabajo es que los genes *alr1976* y *all0345*, identificados como diana directa de FurC, son reguladores transcripcionales que podrían estar implicados en la regulación de metabolismo del nitrógeno, lo que permitiría ampliar el regulón de FurC. En concreto, la expresión del gen *alr1976* disminuye 2,5 veces en *Anabaena* sp. PCC 7120 tras 12 h de cultivo en deficiencia de nitrógeno combinado en comparación con el cultivo en condiciones estándar (67). Por su parte, el gen *all0345* muestra expresión diferencial en la estirpe de sobreexpresión de FurC frente a la estirpe silvestre únicamente en condiciones de deficiencia de nitrógeno, pero no en condiciones estándar (64). Por ello, dado que FurC juega un papel clave en el metabolismo del nitrógeno (64) se plantea que además de regular genes implicados en este proceso de forma directa podrían regular también genes de forma indirecta a través de los reguladores Alr1976 y All0345.

#### 3.2. Objetivos

El propósito de este trabajo Fin de Máster es la purificación y estudio funcional de las proteínas Alr1976 y All0345 de *Anabaena* sp. PCC7120, que están directamente reguladas por FurC y además muestran indicios de participar en el metabolismo del nitrógeno. Para ello, en este trabajo se plantearon los siguientes objetivos:

- Realizar un estudio bioinformático de las proteínas Alr1976 y All0345 para identificar sus dominios, determinar las familias a la que pertenecen, modelar su estructura y estudiar sus relaciones filogenéticas.
- Obtener las proteínas Alr1976 y All0345 mediante sobreexpresión recombinante en *E. coli*, llevando a cabo un clonaje de los genes *alr1976* y *all0345* en el vector pET28a y optimizando su sobreexpresión y posterior purificación mediante cromatografía de afinidad a metales inmovilizados (IMAC).
- Caracterizar bioquímicamente las proteínas Alr1976 y All0345 obtenidas tras la purificación, estudiando la posible oligomerización o presencia de metales y optimizar sus condiciones de unión al DNA mediante ensayos de retardo en gel.
- Estudiar la posible implicación de las proteínas Alr1976 y All0345 en la regulación del metabolismo del nitrógeno, estudiando su unión a regiones promotoras de genes involucrados en el metabolismo del nitrógeno y la diferenciación de heterocistos mediante ensayos de retardo en gel.

## 4. Materiales y métodos

### 4.1 Microorganismos y condiciones de cultivo

#### 4.1.1. Cepas de *E. coli*

Las cepas de *Escherichia coli* utilizadas en este trabajo se recogen en la tabla 1.

Cepa	Genotipo	Empleo	Origen
DH5α	F <sup>-</sup> , U169 (φ80Δ <i>lacZ</i> ΔM15), <i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1 hsdR17</i> ( <i>rK</i> <sup>-</sup> , <i>mK</i> <sup>+</sup> ), <i>phoA</i> , <i>supE44</i> , <i>relA1λ</i> <sup>-</sup> , <i>deoR</i> , Δ( <i>lacZYA-argF</i> )	Conservación de plásmidos	Invitrogen
BL21 (DE3)	F <sup>-</sup> , <i>ompT</i> , <i>hsdS<sub>B</sub></i> ( <i>r<sub>B</sub></i> <sup>-</sup> , <i>m<sub>B</sub></i> <sup>-</sup> ), <i>dcm</i> , <i>gal</i> , λ (DE3)	Sobreexpresión de proteínas	Novagen
DH5α pET <i>alr1976</i>	Células DH5α que contienen el plásmido pET28a: <i>alr1976</i> . Resistente a kanamicina a 50 μg/mL	Conservación del plásmido pET28a: <i>alr1976</i>	Este trabajo
DH5α pET <i>all0345</i>	Células DH5α que contienen el plásmido pET28a: <i>all0345</i> . Resistente a kanamicina a 50 μg/mL	Conservación del plásmido pET28a: <i>all0345</i>	Este trabajo
BL21 pET <i>alr1976</i>	Células BL21 que contienen el plásmido pET28a: <i>alr1976</i> . Resistente a kanamicina a 50 μg/mL	Sobreexpresión de Alr1976	Este trabajo
BL21 pET <i>all0345</i>	Células BL21 que contienen el plásmido pET28a: <i>all0345</i> . Resistente a kanamicina a 50 μg/mL	Sobreexpresión de All0345	Este trabajo

**Tabla 1.** Cepas de *E. coli* utilizadas en el trabajo

#### 4.1.2. Condiciones de cultivo

El cultivo de las cepas de *E. coli* se realizó en medio LB compuesto por 10 g/L de triptona, 5 g/L de extracto de levadura y 10 g/L de NaCl, preparado en agua destilada y esterilizado mediante un autoclave. Los cultivos en medio líquido se incubaron a 37 °C y 200 rpm en un incubador orbital y los cultivos en medio sólido se incubaron a 37 °C en placas de Petri en las que se suplementó el medio LB con agar al 1,5 % (p/v). La selección de las cepas de *E. coli* conteniendo los plásmidos pET28a:*alr1976* y pET28a:*all0345*, tanto en los cultivos líquidos como en cultivos en medio sólido, se llevó a cabo con el antibiótico kanamicina a 50 μg/mL, el cual fue adicionado a partir de un stock 1000 veces más concentrado que estaba disuelto en agua destilada y esterilizado mediante filtración con filtros de 0,22 μm.

## 4.2. Técnicas de trabajo con DNA

#### 4.2.1. Extracción de DNA en *E. coli*

Para verificar que el clonaje de los plásmidos pET28a:*alr1976* y pET28a:*all0345* había tenido lugar de forma correcta, se realizó un análisis por PCR colonia, para lo que fue necesario la extracción de DNA de colonias de células DH5α. Para ello, se tomaron colonias en esterilidad y se resuspendieron en 30 μL de agua Milli-Q en termotubos de 0,2 mL. Los tubos se sometieron a dos ciclos de congelación en nitrógeno líquido seguida de una incubación a 98 °C durante 5 min para poder romper las células y liberar el DNA. Por otro lado, tras el clonaje, los plásmidos pET28a:*alr1976* y pET28a:*all0345* se extrajeron de las células DH5α utilizando el kit GenElute™ Plasmid Miniprep (Sigma-Aldrich) siguiendo las instrucciones del fabricante.

#### 4.2.2. Reacción en cadena de la polimerasa (PCR)

Para llevar a cabo la amplificación tanto de los genes *alr1976* y *all0345* para su posterior clonaje en el vector pET28a como de los promotores para realizar los ensayos de retardo en gel se empleó la reacción en cadena de la polimerasa (PCR). En el caso de la amplificación de los genes *alr1976* y *all0345* para su clonaje en el plásmido pET28a se empleó la polimerasa de alta fidelidad Phusion™, ya que su tasa de error es 50 veces menor que la de la Taq polimerasa. Para llevar a cabo estas reacciones de PCR se preparó una mezcla de reacción compuesta por 25 µL de agua Milli-Q, 9 µL de Buffer HF 5x, 0,9 µL de dNTPs 10 mM, 4,5 µL de cebador directo 4 µM, 4,5 µL de cebador reverso 4 µM, 1 µL de DNA genómico de *Anabaena* sp. PCC 7120 a 100 ng/µL y 0,3 µL de polimerasa Phusion™ (2U/ µL), quedando un volumen total de 45 µL.

Para amplificar las regiones promotoras que se utilizaron en los ensayos de retardo en gel se llevaron a cabo reacciones de PCR en las que se empleó Taq polimerasa y se preparó una mezcla de reacción compuesta por 72 µL de agua Milli-Q, 10 µL de Buffer 10x (Tris-HCl 75 mM pH 9, KCl 50 mM y (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 20 mM), 3 µL de MgCl<sub>2</sub> 50 mM, 2 µL de dNTPs 10 mM, 5 µL de cebador directo 20 µM, 5 µL de cebador reverso 20 µM, 1 µL de Taq DNA polimerasa (Biotools) y 2 µL de DNA genómico de *Anabaena* sp. PCC 7120 a 100 ng/µL, quedando un volumen total de 100 µL.

Además, se llevó a cabo una PCR colonia para verificar que el clonaje en el plásmido pET28a era correcto, para lo que se prepararon mezclas de reacción compuestas por 5,7 µL de agua Milli-Q, 1 µL de Buffer 10x con MgCl<sub>2</sub> (Tris-HCl 75mM pH 9, KCl 50 mM, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 20 mM y MgCl<sub>2</sub> 50 mM), 0,2 µL de dNTPs 10 mM, 1 µL de cebador directo 4 µM, 1 µL de cebador reverso 4 µM, 0,1 µL de Taq Polimerasa 5 U/µL y 1 µL del sobrenadante obtenido tras la extracción de DNA de DH5α.

En todos los casos, la amplificación se llevó a cabo en un termociclador Thermal Cycler 2720 (Applied Biosystems). El programa para PCR y los oligonucleótidos utilizados se recogen en las secciones 9.1, 9.2 y 9.3 de Material Suplementario.

#### 4.2.3. Electroforesis de DNA

La electroforesis de DNA en gel de agarosa se utilizó para comprobar la ausencia de contaminantes y la correcta amplificación de los productos de PCR. Para ello, en primer lugar, se prepararon geles de agarosa al 1% (p/v) con bromuro de etidio al 1 % (v/v), disolviendo 0,5 g de agarosa en 50 mL de TBE (Tris-HCl 90 mM pH 8, ácido bórico 90 mM y EDTA 2 mM) mediante calentamiento en microondas. Esta mezcla se dejó enfriar y se añadieron 50 µL de bromuro de etidio 0,5 g/mL antes de su solidificación. Para llevar a cabo la electroforesis tanto de los promotores como de los genes *alr1976* y *all0345* para su posterior clonaje, se tomaron 4 µL de muestra y se mezclaron con 3 µL de tampón de carga (Tris-HCl 50 mM pH 8, 30 % (v/v) de glicerol y 0,25 % (p/v) de azul de bromofenol) y se cargaron en el gel de agarosa. En el caso de la electroforesis del análisis por la PCR colonia, que se usó para determinar en qué colonias bacterianas el vector pET28a contenía el inserto, se tomaron los 10 µL de muestra, se mezclaron con 3 µL de tampón de carga y se cargaron en el gel de agarosa. Las electroforesis se realizaron en tampón TBE a 90 V durante 30 min, empleando una fuente BioRad PowerPac300 y se visualizaron los resultados en un equipo GelDoc200 (BioRad) con el programa MultiAnalist 1.1 (BioRad Laboratories, 1997).

#### 4.2.3. Purificación y cuantificación de productos de PCR

La purificación de los productos de PCR se llevó a cabo con el kit GFX<sup>TM</sup> PCR DNA and Gel band Purification (GE Healthcare) siguiendo el protocolo del fabricante, para eliminar reactivos sobrantes de la PCR que podían interferir en etapas posteriores. Finalmente, se determinó la concentración de los productos de PCR mediante el equipo NanoVue Plus<sup>TM</sup> (GE Healthcare).

#### 4.2.5. Doble digestión y ligación

Para clonar los genes *alr1976* y *all0345* en el plásmido pET28a se realizó una doble digestión tanto del vector como de los genes, en los que se habían introducido sitios de restricción en los oligonucleótidos al amplificarlos por PCR. La digestión se llevó a cabo con las enzimas de restricción FastDigest<sup>TM</sup> NdeI y Sall (ThermoFisher), mezclando 100 ng del DNA a digerir con 2  $\mu$ L de FastDigest<sup>TM</sup> Buffer (ThermoFisher) y 1  $\mu$ L de cada enzima de restricción, completando hasta 20  $\mu$ L con H<sub>2</sub>O Milli-Q. Las digestiones se incubaron a 37 °C durante 2 horas y, transcurrido ese tiempo, se llevó a cabo la ligación, haciendo uso de los extremos cohesivos generados en la digestión. Para ello, se prepararon mezclas de inserto y vector en una proporción molar de 3:1, se añadieron 2  $\mu$ L de Buffer T4 DNA ligase<sup>TM</sup> (ThermoFisher) y 1  $\mu$ L de enzima T4 ligasa (5 U/ $\mu$ L), completando hasta 20  $\mu$ L con H<sub>2</sub>O Milli-Q y se incubó toda la noche a 16 °C.

### 4.3. Métodos genéticos

#### 4.3.1. Preparación de células de *E. coli* termocompetentes

Para poder realizar la transformación de células de *E. coli*, se requiere que estas se encuentren en un estado de competencia. Por ese motivo, fue necesario preparar células DH5 $\alpha$  y BL21 termocompetentes. Para ello, se inoculó una colonia de cada una de las cepas de *E. coli* en 2 mL de LB estéril y se incubó durante la noche a 37 °C y 200 rpm. Posteriormente, se añadieron los 2 mL de cultivo a 200 mL de LB estéril para refrescar el cultivo y se incubaron a 37 °C y 200 rpm hasta que se alcanzó una densidad óptica a 600 nm de entre 0,3 y 0,4. Una vez alcanzada dicha densidad óptica, los cultivos se incubaron 20 minutos en hielo para detener el crecimiento. Transcurridos los 20 minutos, se dividió el cultivo en 4 tubos falcon de 50 mL, los cuales se centrifugaron durante 15 minutos a 3000 g y 4 °C. A continuación, se retiró el sobrenadante y se resuspendieron los pellets celulares en 10 mL de *buffer* 1 (CaCl<sub>2</sub> 0,13 M, MgCl<sub>2</sub> 0,15 M y acetato de sodio 0,04 M) frío y en esterilidad. Posteriormente, se centrifugó durante 15 minutos a 3000 g y 4 °C, se retiró el sobrenadante y las células se resuspendieron en 4 mL de *buffer* 2 (CaCl<sub>2</sub> 0,13 M y glicerol 15% (v/v)) frío y estéril. Finalmente, se repartieron las células en alícuotas de 200  $\mu$ L en eppendorfs estériles, que se almacenaron a -80 °C.

#### 4.3.2. Transformación de *E. coli* mediante choque térmico

El choque térmico se utilizó para transformar células DH5 $\alpha$  con la mezcla de ligación y para introducir los plásmidos pET28a:*alr1976* y pET28a:*all0345* en células BL21. Las células competentes se descongelaron en hielo durante 10 minutos, se añadieron 100 ng de DNA y se incubaron en hielo durante 10 min. Posteriormente se sometieron a un choque térmico a 42 °C en un termobloque durante 1 minuto y se incubaron 1 minuto en hielo. Transcurrido este tiempo, se añadieron 500  $\mu$ L de medio LB estéril y se incubaron a 37 °C y 200 rpm durante 1 hora. Finalmente, se sembraron las células en una placa de LB agar con kanamicina, pues el plásmido pET28a confiere resistencia a este antibiótico, y se incubaron toda la noche a 37 °C.

#### 4.4. Sobreexpresión de proteínas recombinantes

##### 4.4.1. Prueba de sobreexpresión a pequeña escala

Para conocer las condiciones de mayor sobreexpresión de las proteínas Alr1976 y All0345 se llevó a cabo una prueba de sobreexpresión a pequeña escala en la que se ensayaron diferentes tiempos y temperaturas de inducción (2 h a 37 °C, 4 h a 37 °C, 4 h a 30 °C y toda la noche a 15 °C). En primer lugar, se prepararon preinóculos de BL21-pET*alr1976* y de BL21-pET*all0345* para lo que se inocularon colonias en 2 mL de LB estéril y se incubaron durante la noche a 37 °C y 200 rpm. Al día siguiente se tomaron 100 µL de cultivo y se inocularon en cuatro tubos falcon con de 10 mL de LB estéril y 50 µg/mL de kanamicina, uno para cada condición de inducción. Los cultivos se incubaron a 37 °C y 200 rpm hasta que se alcanzó una densidad óptica a 600 nm de 0,6. Posteriormente se dividió cada cultivo de 10 mL en dos cultivos de 5 mL y en uno de los tubos se indujo la expresión de la proteína mediante adición de 5 µL de IPTG 1 M, quedando a una concentración final de 1 mM. A continuación, el cultivo inducido y el cultivo sin inducir se incubaron a 200 rpm a la temperatura y tiempo de cada una de las condiciones a estudiar.

Para analizar la condición en la que se producía una mayor sobreexpresión se realizó una electroforesis en gel de acrilamida con condiciones desnaturalizantes. Para ello, transcurrido el tiempo de inducción se tomó 1 mL de cada cultivo (tanto inducido como no inducido) y se centrifugó durante 5 min a 12000 rpm en una centrífuga Eppendorf 5427R. Posteriormente se descartó el sobrenadante y el precipitado se congeló a -20 °C hasta su uso. Una vez que se obtuvieron todas las muestras, los precipitados se resuspendieron en 75 µL de agua Milli-Q y 15 µL de tampón de carga 6x (60 mM de Tris-HCl pH 6,8, 9,6% (p/v) de SDS, 42 % de glicerol 0,06 % (p/v) de azul de bromofenol y 30 % (v/v) de β- mercaptoetanol). Se hirvieron las muestras a 95 °C durante 5 min, se centrifugaron 5 min a 12000 rpm en una centrífuga Eppendorf 5427R y se cargaron 10 µL de sobrenadante en geles de acrilamida al 17 %. La electroforesis se llevó a cabo como se detalla en el apartado 4.6.2.

##### 4.4.2. Prueba de solubilización

Antes de llevar a cabo la purificación, se estudió cuál era el tampón óptimo para la solubilización de la proteína, para ello se probaron cuatro tampones diferentes. El primero estaba compuesto por Tris 50 mM pH 8, 10 % de glicerol y 500 mM de NaCl; el segundo por Tris 50 mM pH 8, NaCl 200 mM, 10 % de glicerol y Tween 0,05 %; el tercero por Tris 50 mM, pH 8 y 1 U de lisozima y el último por NaH<sub>2</sub>PO<sub>4</sub> 0,1 M, Tris 0,01 M y Cloruro de guanidinio 2 M a pH 8.

Para llevar a cabo la prueba de solubilización se descongelaron 50 mg de biomasa de células BL21 sobreexpresando la proteína Alr1976 o All0345 y se resuspendieron en 1 mL del tampón a analizar. Se añadieron 10 µL de PMSF 100 mM (inhibidor de proteasas), de modo que la concentración final fue de 1 mM y en el tampón que contenía lisozima se incubó 30 min a 37 °C. Para romper las células se emplearon 10 ciclos de sonicación de 45 s a una amplitud del 80 % y ciclo 0,5, con 30 s de reposo entre los ciclos y tras la sonicación, las muestras se centrifugaron a 4 °C, 14000 rpm durante 15 min en una centrífuga Eppendorf 5427R para separar la fase soluble (sobrenadante) de la fase insoluble (precipitado).



Para estudiar la condición que proporcionaba una mejor solubilidad de la proteína, se realizó una electroforesis desnaturizante en gel de poliacrilamida. Para ello, se tomaron 16  $\mu$ L de sobrenadante, los cuales fueron sometidos a una precipitación con ácido tricloroacético (TCA) al 10 %, que fue llevada a cabo como se detalla en el apartado 4.6.1 antes de ser cargados en el gel de electroforesis. También se analizó la fase insoluble para conocer cuanta proteína no era soluble, para lo cual se resuspendió el precipitado en 200  $\mu$ L de tampón de carga 6x (60 mM de Tris-HCl pH 6,8, 9,6% (p/v) de SDS, 42 % de glicerol 0,06 % (p/v) de azul de bromofenol y 30 % (v/v) de  $\beta$ - mercaptoetanol) y 800  $\mu$ L de agua Milli-Q. Finalmente, se llevó a cabo la electroforesis como se describe en el apartado 4.6.2.

### 4.5. Purificación de proteínas recombinantes

#### 4.5.1. Obtención de biomasa

Tras determinar las condiciones de tiempo y temperatura de inducción que permitían una mayor sobreexpresión de proteína, se obtuvo biomasa a gran escala de células BL21-pET*alr1976* y BL21-pET*all0345* sobreexpresando la proteína de interés. Para llevarlo a cabo, en primer lugar, se pusieron preinóculos de células BL21-pET*alr1976* y BL21-pET*all0345* en 10 mL de LB con 50  $\mu$ g/mL de kanamicina, los cuales se incubaron durante toda la noche a 37 °C y 200 rpm. Transcurrido ese tiempo, se refrescaron los cultivos en erlenmeyers con 1 L de medio LB con 50  $\mu$ g/mL de kanamicina y se incubaron a 200 rpm y 37 °C hasta que la densidad óptica a 600 nm fue de 0,6. Una vez alcanzada dicha densidad óptica, se indujo la sobreexpresión de la proteína mediante la adición de 1 mL de IPTG 1M, de forma que la concentración final de IPTG fue de 1 mM y se incubaron a 200 rpm y 15 °C y toda la noche. Transcurrido ese tiempo, los cultivos se centrifugaron a 4 °C y 5000 rpm durante 10 min en una centrífuga Gyrozen 2236R con un rotor GRF-L-500-6. Tras la centrifugación, se descartó el sobrenadante y se guardó el pellet a -20 °C hasta su empleo.

#### 4.5.2. Lisis celular

Para poder purificar las proteínas, primero es necesario lisar las células y que estas liberen su contenido. Para ello se descongelaron 10 g de biomasa y se resuspendieron en 100 mL de tampón de purificación, compuesto por Tris 50 mM, Tween 0,1 % (v/v), glicerol 10 % (v/v), NaCl 200 mM e imidazol 20 mM a pH 8, cuando se realizó la purificación con Tris 50 mM, y cuando la purificación fue realizada con cloruro de guanidinio se resuspendió en 50 mL de tampón de purificación formado por NaH<sub>2</sub>PO<sub>4</sub> 0,1 M, Tris 0,01 M y cloruro de guanidinio 2 M a pH 8. Posteriormente, se añadió PMSF 100 mM, de forma que la concentración final fuese de 1 mM y se lisaron las células mediante 10 ciclos de sonicación de 45 s de amplitud del 80 % y ciclo 0'5 con 30 s de reposo entre ciclos en un sonicador UP 200 S Dr. Hielscher®, manteniendo la muestras en baño de hielo para evitar el calentamiento de la muestra. Las células sonicadas fueron centrifugadas en una centrífuga Gyrozen 2236R con un rotor GRF-L-50-6 a 20000 rpm y 4 °C durante 15 min. El sobrenadante obtenido en la centrifugación se centrifugó otras dos veces en las mismas condiciones para obtener el extracto crudo, que se conservó en frío hasta la purificación de las proteínas.

#### 4.5.3. Purificación por cromatografía de afinidad a iones metálicos inmovilizados (IMAC)

Para purificar las proteínas All0345 y Alr1976 a partir del extracto crudo se llevó a cabo una cromatografía de afinidad a iones metálicos inmovilizados (IMAC, immobilized metal affinity chromatography) con la matriz Chelating Sheparose Fast Flow, ya que los genes *all0345* y *alr1976* se clonaron en el plásmido pET28a con una cola de histidinas en el extremo amino terminal, la cual presenta afinidad por metales divalentes.

Para llevar a cabo la purificación, en primer lugar, se añadió la matriz sobre la columna, se empaquetó por gravedad y posteriormente se le aplicó agua Milli-Q a una velocidad de 3 mL/min, la cual se mantuvo constante durante todo el protocolo de purificación. A continuación, se cargó la columna con el metal, para lo que se aplicaron 3 volúmenes de NiSO<sub>4</sub> 0,3 M o 3 volúmenes de CoCl<sub>2</sub> 0,3 M, en función del metal con el que se desease realizar la purificación. Se eliminó el exceso de metal no adherido con 10 volúmenes de agua Milli-Q y después, para equilibrar la columna, se aplicaron 10 volúmenes de tampón de purificación (Tris 50 mM, Glicerol 10 %, NaCl 200 mM, imidazol 20 mM, a pH 8) en el caso de la purificación con Tris 50 mM pH 8 y con 10 volúmenes de una solución de KCl 0,4 M, NaH<sub>2</sub>PO<sub>4</sub> 0,1 M, Tris 0,01 M y Cloruro de guanidinio 2 M, a pH 8 en el caso de la purificación con cloruro de guanidinio. Una vez añadidos estos volúmenes, se cargó el extracto proteico y el volumen eluido de la columna (*flowthrough*) se recogió para verificar que la proteína se había fijado a la columna.

Tras aplicar el extracto crudo, se lavó la columna para eliminar todas aquellas proteínas que se hubieran unido a la columna por interacciones inespecíficas con 15 volúmenes de tampón de lavado (tampón de purificación con 40 mM de imidazol) en el caso de la purificación con Tris 50 mM y con 5 volúmenes de una solución de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 0,5 M disuelta en tampón de purificación ((NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 0,5 M, NaH<sub>2</sub>PO<sub>4</sub> 0,1 M, Tris 0,01 M y cloruro de guanidinio 2 M, a pH 8) seguidos de 10 volúmenes de una solución de glicina 35 mM en tampón de purificación en el caso de la purificación con cloruro de guanidinio. En ambos casos se comprobó que la densidad óptica del eluido a 280 nm era menor de 0,1. Finalmente, las proteínas de interés se eluyeron con un gradiente creciente de imidazol (0-1 M) disuelto en el tampón de purificación correspondiente para cada purificación y el eluido se recogió en alícuotas de 1 mL, las cuales fueron estudiadas mediante una electroforesis desnaturizante en gel de poliacrilamida (apartado 4.6.2), previa precipitación con ácido tricloroacético (apartado 4.6.1).

Se seleccionaron las alícuotas en las que la proteína de interés estuviera en mayor cantidad y con menos contaminantes, que se combinaron y sometieron a diálisis, para poder eliminar el imidazol empleado en la purificación. Se realizaron dos ciclos de diálisis a 4 °C en 3 L de tampón Tris 50 mM, glicerol 10 % (v/v) y NaCl 100 mM, a pH 8 en el caso de la purificación realizada con Tris 50 mM pH 8 y en 3 L de tampón acético-acetato 10 mM y pH 4 en el caso de la purificación con cloruro de guanidinio. Finalmente, las proteínas fueron cuantificadas por espectroscopía, como se detalla en el apartado 4.6.3.

Finalmente se regeneró la columna pasando 5 volúmenes de agua Milli-Q, 5 volúmenes de *Strip buffer* (EDTA 50 mM, NaCl 500 mM, a pH 8), 3 volúmenes de NaCl 0,5 M, 5 volúmenes de agua Milli-Q y finalmente 5 volúmenes de etanol al 20 % (v/v) para preservar la resina.

## 4.6. Técnicas analíticas de proteínas

### 4.6.1. Precipitación con ácido tricloroacético (TCA)

Dado que tanto el imidazol como el cloruro de guanidinio pueden interferir con la electroforesis, para poder analizar mediante electroforesis desnaturalizante las fracciones obtenidas tras eluir las columnas en la purificación, así como para determinar el tampón que proporcionaba una mejor solubilidad, se realizó una precipitación previa con ácido tricloroacético (TCA). Para ello, se mezclaron 16  $\mu\text{L}$  de sobrenadante en el caso del ensayo de sobreexpresión o de alícuota en el caso de la purificación con 84  $\mu\text{L}$  de agua Milli-Q y 100  $\mu\text{L}$  de TCA al 10 % (v/v) y se incubaron durante 15 min en hielo. Transcurrido ese tiempo, se centrifugó a 12000 rpm y 4 °C durante 10 min en una centrífuga Eppendorf 5427R. Tras la centrifugación, se descartó el sobrenadante y se añadieron 500  $\mu\text{L}$  de acetona fría. Posteriormente, se centrifugó a 4 °C y 12000 rpm durante 10 min en una centrífuga Eppendorf 5427R, se descartó el sobrenadante y se dejó secar el precipitado durante 10 min a temperatura ambiente, para eliminar la acetona restante. Finalmente, los precipitados se resuspendieron en 16  $\mu\text{L}$  de agua Milli-Q.

### 4.6.2. Electroforesis desnaturalizante en geles de poliacrilamida

Para poder analizar los resultados de los ensayos de sobreexpresión y solubilidad, las diferentes etapas de la purificación y realizar los ensayos de oligomerización, se llevó a cabo una electroforesis desnaturalizante en geles de poliacrilamida con SDS (SDS-PAGE). Los geles estaban formados por una región concentradora con un 5 % de acrilamida/bisacrilamida y una región separadora con un 15-17 % de acrilamida/bisacrilamida, que se prepararon en un sistema Dual Gel Caster (Bio-Rad). La composición de los geles de acrilamida/bisacrilamida se encuentra detallada en la sección 9.4 de Material Suplementario.

Las muestras se prepararon mezclando 16  $\mu\text{L}$  de muestra con 4  $\mu\text{L}$  de tampón de carga 6x (60 mM de Tris-HCl pH 6,8, 9,6% (p/v) de SDS, 42 % de glicerol 0,06 % (p/v) de azul de bromofenol y 30 % (v/v) de  $\beta$ - mercaptoetanol), que se hirvieron durante 5 minutos a 95 °C para desnaturalizar las proteínas y se sometieron a un pulso de centrifuga. Se cargaron 20  $\mu\text{L}$  de muestra tras haber limpiado los pocillos del gel manualmente. La electroforesis se llevó a cabo con una corriente constante de 35 mA por gel durante 45 minutos en tampón Tris-HCl 25 mM, pH 8,8, glicina 200 mM y SDS 3,5 mM empelando un dispositivo de electroforesis vertical Mini-PROTEAN® Tetra Vertical Electrophoresis Cel (Bio-Rad) y una fuente de corriente PowerPac Basic (Bio-Rad).

Para poder visualizar las proteínas, los geles de poliacrilamida se tiñeron con una solución colorante compuesta por ácido acético al 10 % (v/v), metanol al 45 % (v/v), agua al 45 % (v/v) y Coomassie Blue R-250 al 0,25 % (p/v) durante al menos 20 minutos. Para eliminar el exceso de azul de Coomassie, se realizaron lavados sucesivos con una solución decolorante compuesta por ácido acético al 10 % (v/v), metanol al 25 % (v/v) y agua al 65 % (v/v).

### 4.6.3. Cuantificación de las proteínas por espectroscopía

Para cuantificar las proteínas se realizó un espectro en la región UV-visible de las soluciones protéicas. La concentración de las proteínas se calculó mediante la ley de Beer-Lambert ( $Abs = \epsilon \cdot c \cdot l$ ) en la que *Abs* corresponde a la Absorbancia a 280 nm (longitud de onda en la que absorben los aminoácidos aromáticos de la proteína);  $\epsilon$ , al coeficiente de extinción molar teórico a 280 nm; *c*, a la concentración molar y *l*, a la longitud de paso de la cubeta (en este caso, 1 cm).



## 4.7. Caracterización bioquímica de proteínas

### 4.7.1. Determinación de la presencia de metales por ICP-OES

Para determinar la presencia de metales en las soluciones de proteína, se llevó a cabo un análisis mediante ICP-OES (Inductively Coupled Plasma – Optical Emission Spectroscopy), utilizando como control una muestra de tampón acético-acetato 10 mM pH 4, en el cual se encontraban las proteínas. Para preparar las muestras que se analizaron mediante ICP-OES, se añadieron 3 mL de ácido nítrico al 5% para ICP en un tubo falcón de 15 mL, se añadió el volumen de solución de proteína para tener un total de 10 nmol de proteína y finalmente se enrasó con ácido nítrico al 5% para ICP hasta los 10 mL. Tras obtener los resultados y sabiendo la cantidad de proteína que se añadió, se calculó la cantidad de metal por mol de proteína.

### 4.7.2. Ensayos de retardo en gel (EMSA)

Para determinar si las proteínas All0345 y Alr1976 purificadas eran activas y presentaban actividad de unión a DNA, así como para estudiar su unión *in vitro* a distintos promotores, se realizaron ensayos de retardo en gel (*Electrophoretic Mobility Shift Assay*, EMSA). Este ensayo se basa en la capacidad de los reguladores transcripcionales de unirse *in vitro* al DNA diana, generando un complejo DNA-proteína, el cual presenta una migración retardada respecto al DNA libre en una electroforesis en condiciones nativas. Los ensayos se realizaron en presencia de un DNA competidor, en este caso un fragmento interno del gen *pkn22* (*ifpkn22*) para demostrar la unión específica de la proteína por el DNA de interés.

Los ensayos de retardo en gel se realizaron en geles de poliacrilamida no desnaturizantes al 6 %, cuya composición se recoge en la sección 9.5 del Material Suplementario, que es un grado de entrecruzamiento que permite la migración de los complejos DNA-proteína. Dichos geles se prepararon en un sistema Dual Gel Caster (BioRad) y se precorrieron durante al menos 30 min a 60 V y 4 °C en tampón de electroforesis compuesto por Tris-HCl 50 mM y Glicina 380 mM a pH 8,5, con el objetivo de eliminar restos de poliacrilamida y persulfato de amonio de los pocillos, equilibrar el gel en el tampón de electroforesis y conseguir una temperatura constante.

Se prepararon mezclas de reacción en un volumen final de 20 µL que estaban compuestas por 50 ng de la región promotora de interés (su propio promotor para estudiar las condiciones de unión al DNA de cada proteína y promotores de genes relacionados en el metabolismo del nitrógeno en el caso de la búsqueda de dianas), 50 ng de DNA competidor (*ifpkn22*), 2 µL de tampón de unión (Bis-Tris 10 mM pH 7,5, KCl 40 mM, glicerol 5 % (v/v)), 1 µL de albúmina sérica bovina a 1 mg/mL, completando el volumen con agua Milli-Q. Las mezclas se incubaron durante 30 min a temperatura ambiente y, transcurrido ese tiempo, se añadieron 3 µL de tampón de carga (Tris-HCl 50 mM pH 8, glicerol 30 % (v/v) y azul de bromofenol 0,05 % (p/v)). Posteriormente, se limpiaron los pocillos del gel y se cargaron 20 µL de muestra, incluyendo un control negativo sin proteína. La electroforesis se realizó a 90 V, a 4 °C y durante 90 min en tampón de electroforesis (Tris-HCl 50 mM y Glicina 380 mM a pH 8,5). Finalmente, se tiñeron los geles en cubetas con 50 mL de tampón de electroforesis y 3 µL de SyBr® Safe Fluorescent Sye (Invitrogen) durante 10 min en agitación y oscuridad y se revelaron en un transiluminador Gel Doc 2000.

### 4.7.3. Dicroísmo circular

El dicroísmo circular permite obtener información sobre la estructura secundaria de la proteína si las medidas se realizan en la región lejana del ultravioleta y sobre la estructura terciaria si se realizan en la región cercana. Los espectros de dicroísmo se obtuvieron en un espectrofotómetro Chirscan empleando soluciones de proteína a una concentración de 10  $\mu\text{M}$ . Las medidas realizadas en el ultravioleta lejano se tomaron a longitudes de onda de entre 190 y 260 nm, empleando cubetas de paso de luz de 0,1 cm con 300  $\mu\text{L}$  de muestra y las realizadas en el ultravioleta se tomaron a longitudes de onda de entre 250 y 310 nm, empleando cubetas de paso de luz de 1 cm con 1000  $\mu\text{L}$  de muestra. En ambos casos, se tomaron medidas cada 0,5 nm del intervalo de longitud de onda seleccionada y se realizaron 8 mediciones por punto, obteniendo al final la media en cada longitud de onda estudiada. Los resultados se expresaron como elipticidad media de los residuos (MRE) en  $\text{deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$  definida como

$$MRE = \frac{\theta_{obs}(mdeg) \cdot M}{n \cdot l \cdot C}$$

Donde  $\theta_{obs}$  es la elipticidad en mdeg,  $M$  es el peso molecular en  $\text{g} \cdot \text{dmol}^{-1}$ ,  $n$  es el número de aminoácidos,  $l$  es la longitud de paso de la cubeta y  $C$  es la concentración de proteínas en  $\text{g} \cdot \text{L}^{-1}$ . Los espectros se deconvolucionaron empleando el servidor BeStSel (<https://bestsel.elte.hu/index.php>).

### 4.7.4. Análisis de oligomerización por entrecruzamiento

Se llevó a cabo un análisis de oligomerización de las proteínas Alr1976 y All0345 mediante entrecruzamiento (*cross-linking*). Esta técnica se basa en el empleo de reactivos bifuncionales que permiten la unión de dos moléculas que, debido a interacciones proteína-proteína, estén muy próximas entre sí. Para realizar este estudio se empleó como agente entrecruzante el glutaraldehído, que permite entrecruzar dos grupos aminos de proteínas próximas.

La reacción se llevó a cabo en un volumen final de 20  $\mu\text{L}$ . En primer lugar, se añadió la cantidad de solución de proteína necesaria para que la concentración final fuera de 10  $\mu\text{M}$  y se incubó 10 min a temperatura ambiente. Posteriormente, se añadió glutaraldehído a una concentración final del 5% (p/v) y se incubaron las muestras a temperatura ambiente durante 25 minutos y en oscuridad, ya que el glutaraldehído es fotosensible. Transcurrido ese tiempo, se añadieron 4  $\mu\text{L}$  de Tampón de carga 6x (60 mM de Tris-HCl pH 6,8, 9,6% (p/v) de SDS, 42 % de glicerol 0,06 % (p/v) de azul de bromofenol y 30 % (v/v) de  $\beta$ - mercaptoetanol) y los resultados se analizaron mediante una electroforesis desnaturante en geles de poliacrilamida formados por una región separadora al 15 % de acrilamida/bisacrilamida y una concentradora al 5% de acrilamida/bisacrilamida como se detalla en el apartado 4.6.2.

### 4.7.5. Ensayos de oligomerización mediada por puentes disulfuro

Con el objetivo de estudiar si las proteínas Alr1976 y All0345 tienen estados de oligomerización mediados por puentes disulfuro se realizó un ensayo de oligomerización, en el que se sometió a las proteínas a condiciones reductoras y oxidantes. Como agente oxidante se empleó el peróxido de hidrógeno ( $\text{H}_2\text{O}_2$ ) y como agente reductor, el 1,4-ditiotreitol (DTT). Las reacciones se prepararon en un volumen final de 20  $\mu\text{L}$  con una concentración de proteína de 10  $\mu\text{M}$  y las muestras se incubaron a temperatura ambiente con DTT 10 mM durante 10 min, con  $\text{H}_2\text{O}_2$  10 mM durante 10 min y primero con  $\text{H}_2\text{O}_2$  10 mM durante 10 min y después con DTT 10 mM durante 10 min, incluyendo también un control de proteína sin tratar.

Tras la incubación con los agentes oxidantes o reductores se añadieron 4  $\mu$ L de tampón de carga 6x sin  $\beta$ - mercaptoetanol (60 mM de Tris-HCl pH 6,8, 9,6% (p/v) de SDS, 42 % de glicerol 0,06 % (p/v) de azul de bromofenol) y los resultados se analizaron mediante electroforesis desnaturante en geles de poliacrilamida formados por una región separadora al 15 % de acrilamida/bisacrilamida y una región concentradora al 5% de acrilamida/bisacrilamida (apartado 4.6.2.).

#### 4.8. Técnicas bioinformáticas

Para conocer las características fisicoquímicas teóricas de las proteínas Alr1976 y All0345 como el número de aminoácidos, peso molecular, punto isoeléctrico (pI) y coeficiente de extinción, las cuales son de utilidad a la hora de abordar su purificación, se empleó la herramienta ProtParam de Expasy (<https://web.expasy.org/protparam/>). Para ello, se utilizaron las secuencias de aminoácidos de las proteínas de la base de datos KEGG Genes (<https://www.genome.jp/kegg/>). Además, con el objetivo de determinar la familia, dominios y motivos que poseían las proteínas de estudio se emplearon los servidores Prosite (<https://prosite.expasy.org/>), Interpro (<https://www.ebi.ac.uk/interpro/>) y SMART (<http://smart.embl-heidelberg.de/>)

La predicción de estructura secundaria de las proteínas se llevó a cabo empleando el servidor PSIPRED (<http://bioinf.cs.ucl.ac.uk/psipred>) y los modelos de estructura terciaria se obtuvieron empleando la base de datos AlphaFold (<https://alphafold.com/>), que lleva a cabo un modelado empleando redes neuronales y el servidor Swiss-Model (<https://swissmodel.expasy.org/>). Las superposiciones de estructuras terciarias se obtuvieron mediante la herramienta ‘MatchMaker’ de Chimera, con los parámetros por defecto. Los modelos de las proteínas unidas al DNA se construyeron alineando los modelos construidos sobre los complejos proteína-DNA que habían sido modelados o resueltos de forma experimental para las proteínas empleadas como molde mediante la herramienta ‘MatchMaker’ de Chimera. Todas las estructuras se visualizaron en UCSF Chimera (<https://www.cgl.ucsf.edu/chimera/download.html>)

Con el objetivo de estudiar la conservación de estas proteínas en otras cianobacterias se buscaron homólogos en cianobacterias mediante la herramienta BLASTP de NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), seleccionando secuencias de diferentes géneros cianobacterianos. También se estudio su similitud con reguladores de la misma familia en otros procariotas, para lo cual se buscaron en Uniprot (<https://www.uniprot.org/>) proteínas en otros microorganismos que fueran de la misma familia que las proteínas de estudio.

En ambos casos, se construyo un alineamiento múltiple mediante la herramienta Clustal W ([http://www.phylogeny.fr/one\\_task.cgi?task\\_type=clustalw](http://www.phylogeny.fr/one_task.cgi?task_type=clustalw)) y un árbol filogenético mediante PhyML ([http://www.phylogeny.fr/one\\_task.cgi?task\\_type=phyml](http://www.phylogeny.fr/one_task.cgi?task_type=phyml)) y la herramienta Phylogeny (Phylogeny.fr: "A la Carte" Mode), que se visualizó con TreeDyn ([http://www.phylogeny.fr/one\\_task.cgi?task\\_type=treedyn](http://www.phylogeny.fr/one_task.cgi?task_type=treedyn)). Para obtener el soporte estadístico de las ramas se utilizaron 100 bootstraps en el cálculo del árbol. Los alineamientos múltiples se visualizaron en JalView (<https://www.jalview.org/>).

## 5. Resultados

### 5.1. Análisis bioinformático de los reguladores Alr1976 y All0345

Antes de abordar la purificación de los reguladores Alr1976 y All0345 se llevó a cabo un análisis bioinformático para obtener información de interés sobre estas proteínas. En primer lugar, se obtuvo su secuencia de aminoácidos en KEGG Genes y mediante ProtParam se determinaron sus características fisicoquímicas (Tabla 2), las cuales son de utilidad para diseñar un protocolo de purificación. Por una parte, se requiere conocer el punto isoeléctrico (pI) para poder solubilizarlas en el pH adecuado y de esa forma evitar la precipitación de las proteínas, y el coeficiente de extinción molar ( $\epsilon$ ) para así poder cuantificar la concentración obtenida tras la purificación mediante espectroscopía. Por otro lado, también es necesario conocer el peso molecular para poder identificar las proteínas en los geles de electroforesis desnaturalizante.

	<b>Alr1976</b>	<b>All0345</b>
<b>Nº Aminoácidos</b>	121	290
<b>Peso molecular (Da)</b>	13456	32946
<b>pI</b>	4,88	5,72
<b>Coeficiente de extinción (<math>M^{-1}cm^{-1}</math>)</b>	8480	37945

**Tabla 2.** Características biofísicas teóricas de las proteínas Alr1976 y All0345 obtenidas mediante ProtParam de ExPasy (<https://web.expasy.org/protparam/>) a partir de su secuencia de aminoácidos.

Dado que en Uniprot ambas proteínas estaban únicamente anotadas como reguladores transcripcionales, se realizó una búsqueda de las proteínas en Prosite, Interpro y SMART para conocer a que familia de reguladores pertenecen, así como la presencia de dominios característicos, cuyo resultado se muestra en la Figura 5.

- A.** MGSYLDIERLAS **LVRKKRGSRLRETSVEIGNVSPSTLSRVESGRMPDMETFLALCNWLQVPPAELF**  
RTTEEDQPDTPAIAIQLRADKKLDPAIANALASLVKAAYRDL SQQDDELKQDP
- B.** MYCRQLDRLSNPYHLQRSKML **LKIGDFSKLSQVSVKALRLYDQMGLLKPIKVDNFTSYRYSSAEQLP**  
**RLNRILAFKDLGFSLEQIAKLLDEN**LPPEQIRGMLRLKQGEIQRLVEAEQARLIRVEARLKQIEQED  
SMPN **YEVVIKKVAPIQVASIRQILPDNPSIGQLYGEISEYLAQNGVKAGDYAGIWHDPGYKDTDID**  
**AEAVISIEGSIKGNRIKIYELPGSETTACLIHHGSYETLAQAYATLVSWIEANGYNITAPNREVYI**  
**IGGNEQNND SYVTELQFPVAQA**

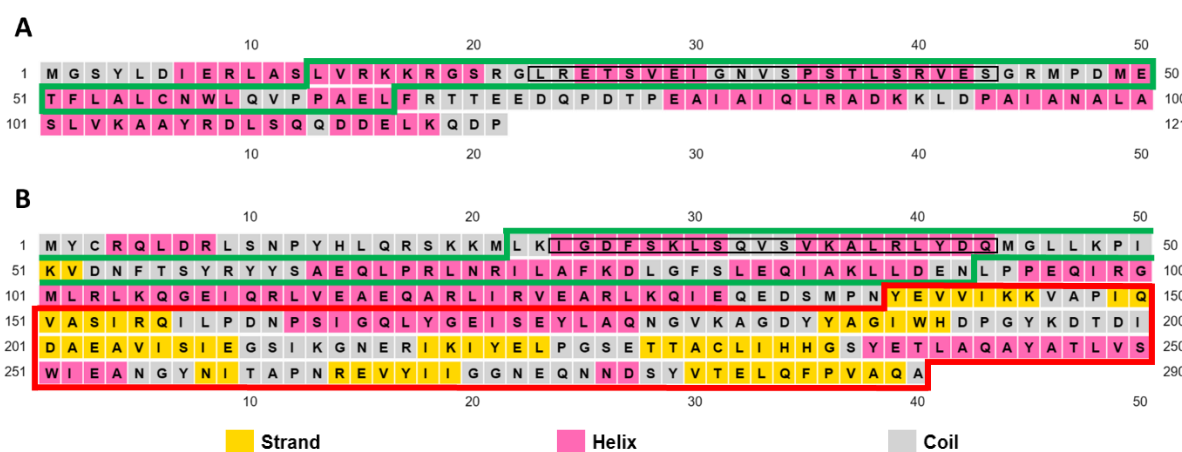
**Figura 5. Identificación de dominios en las proteínas Alr1976 y All0345.** **A.** Secuencia de aminoácidos de la proteína Alr1976, en la que se indica en verde el dominio HTH de unión a DNA de tipo XRE, dentro del cual se resalta en negrita el motivo HTH de unión a DNA. **B.** Secuencia de aminoácidos de la proteína All0345, en la que se indica en verde el dominio HTH de unión a DNA de tipo MerR, dentro del cual se resalta en negrita el motivo HTH de unión a DNA, y en rojo el dominio de unión a efector tipo AraC-E.

Se encontró que la proteína Alr1976 es un regulador transcripcional de la familia XRE (Xenobiotic Response Element) que contiene un dominio HTH (hélice-giro-hélice) de unión a DNA de tipo XRE (también denominado de tipo Cro/C1 o de tipo represor lambda) comprendido entre los aminoácidos 13 a 66, con un motivo HTH de unión a DNA entre los residuos 23 y 43 (Figura 5A). Respecto a la proteína All0345, se determinó que es un regulador transcripcional de la familia MerR y que también posee un dominio HTH de unión a DNA, en este caso de tipo MerR, que abarca de los aminoácidos 22 a 92 e incluye un motivo HTH de unión a DNA entre los residuos 24 y 43. Además, a diferencia de Alr1976, el regulador All0345 también posee en el extremo C-terminal un dominio de unión a efector de tipo AraC-E (o dominio de unión GyrL) entre los residuos 139 y 288 (Figura 5B).

La familia de reguladores XRE se encuentra ampliamente distribuida en bacterias, así como en arqueas y eucariotas. Los miembros de esta familia funcionan como represores en condiciones estándar y generalmente se unen a sus promotores diana en forma de dímeros (68). Estos reguladores comparten un dominio N-terminal de unión al DNA con un motivo HTH, mientras que la región C-terminal es altamente variable (69). El dominio de unión a DNA está compuesto por 5 hélices  $\alpha$ , constituyendo la segunda y la tercera el motivo HTH (68).

Por otro lado, los miembros de la familia MerR, a la que pertenece AlI0345, se caracterizan por poseer en su extremo N-terminal, el dominio de unión a DNA de tipo MerR (el cual está altamente conservado y que posee el motivo HTH), una hélice central que permite su dimerización y un dominio C-terminal de unión a efector (70). En función de su mecanismo de regulación, los miembros de la familia MerR se suelen dividir en tres subfamilias: el primer grupo está formado por los miembros que responden a metales, el segundo, por los que responden a estrés oxidativo y el tercero por los que responden a metabolitos (71), pudiendo detectar múltiples compuestos con propiedades químicas diversas. Los reguladores de esta familia se suelen unir a los promotores de genes diana en forma de dímeros y generalmente actúan como activadores (72).

Para obtener información sobre el plegamiento de las proteínas de interés se realizaron predicciones tanto de su estructura secundaria como de su estructura terciaria. En primer lugar, se predijo la estructura secundaria empleando el servidor PSIPred (Figura 6).

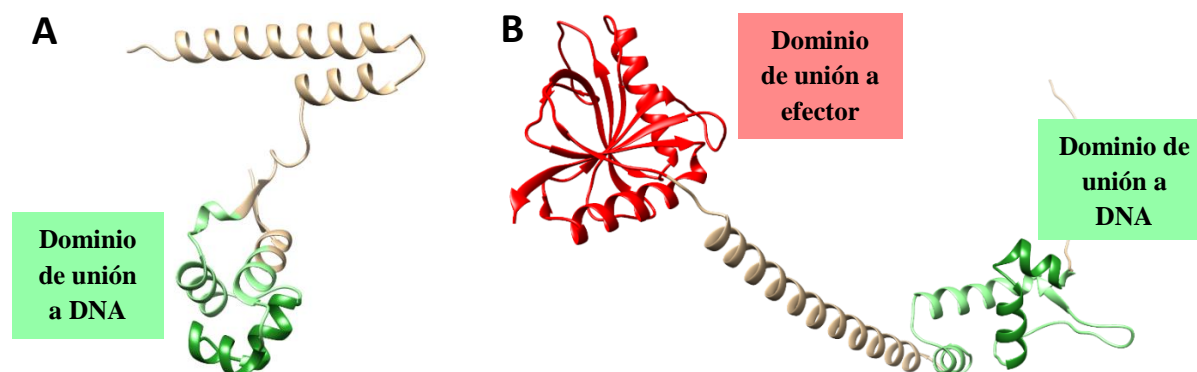


**Figura 6. Predicción de la estructura secundaria de las proteínas Alr1976 (A) y AlI0345 (B) mediante PSIPRED.** En ambos casos aparece recuadrado en verde el dominio de unión a DNA, con el motivo HTH recuadrado en negro. En el caso de AlI0345 se recuadra además en rojo el motivo de unión a efector AraC-E.

Como se muestra en la figura 6A, se predice que la proteína Alr1976 se compone de ocho hélices  $\alpha$ . Las cinco primeras hélices forman parte del dominio de unión a DNA de tipo XRE, quedando tres hélices en la región carboxilo terminal, aunque el bucle que separa las dos últimas hélices solo tiene un residuo y por tanto es posible que ambas formen una única hélice. Con respecto a la proteína AlI0345, se predice que se compone de nueve hélices  $\alpha$  y diez láminas  $\beta$  (figura 6B). Se observa que el dominio de unión a DNA de tipo MerR se compone de cuatro hélices  $\alpha$  y una pequeña lámina  $\beta$  y que el dominio efector de tipo AraC-E se compone tanto de hélices  $\alpha$  como láminas  $\beta$ , estando ambos dominios conectados por una larga hélice  $\alpha$ . Además, se observa que en ambos casos los motivos HTH (hélice-giro-hélice) de unión a DNA se componen, como su nombre indica, de dos hélices separadas por un bucle.



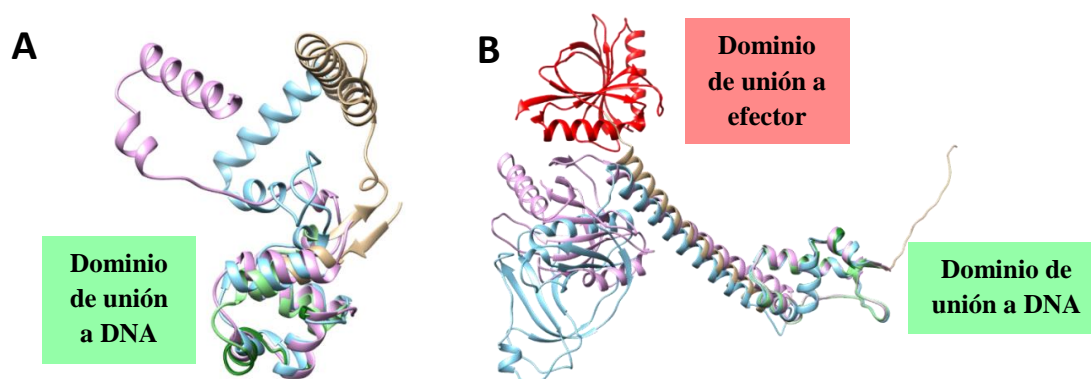
A continuación, se llevó a cabo una predicción de estructura terciaria empleando el servidor AlphaFold, cuyo resultado se muestra en la Figura 7.



**Figura 7. Predicción de la estructura terciaria de las proteínas Alr1976 (A) y All0345 (B) mediante el servidor AlphaFold.** Se representan las estructuras terciarias con un diagrama de cintas. En ambas proteínas se indica en color verde el dominio de unión a DNA, con el motivo HTH resaltado en verde oscuro y en la proteína All0345 se indica además en rojo el dominio de unión a efector de tipo AraC-E.

En la figura 7A se puede observar que la estructura predicha para Alr1976 presenta cinco hélices  $\alpha$  en el dominio de unión a DNA de tipo XRE, una pequeña hoja  $\beta$  formada por dos láminas  $\beta$  que se encuentran al comienzo y al final del dominio de unión al DNA y dos hélices  $\alpha$  en el extremo carboxilo terminal. Por otro lado, la proteína All0345 contiene cuatro hélices  $\alpha$  y una lámina  $\beta$  en el dominio de unión a DNA de tipo MerR, dos hélices  $\alpha$  y una hoja  $\beta$  antiparalela formada por ocho láminas  $\beta$  en el dominio de unión a efector de tipo AraC-E y una larga hélice  $\alpha$  conectando ambos dominios (Figura 7B).

Además de predecir la estructura terciaria con AlphaFold se llevó a cabo un modelado por homología de ambas proteínas mediante SwissModel, cuyo resultado se comparó con las estructuras predichas por AlphaFold (Figura 8).

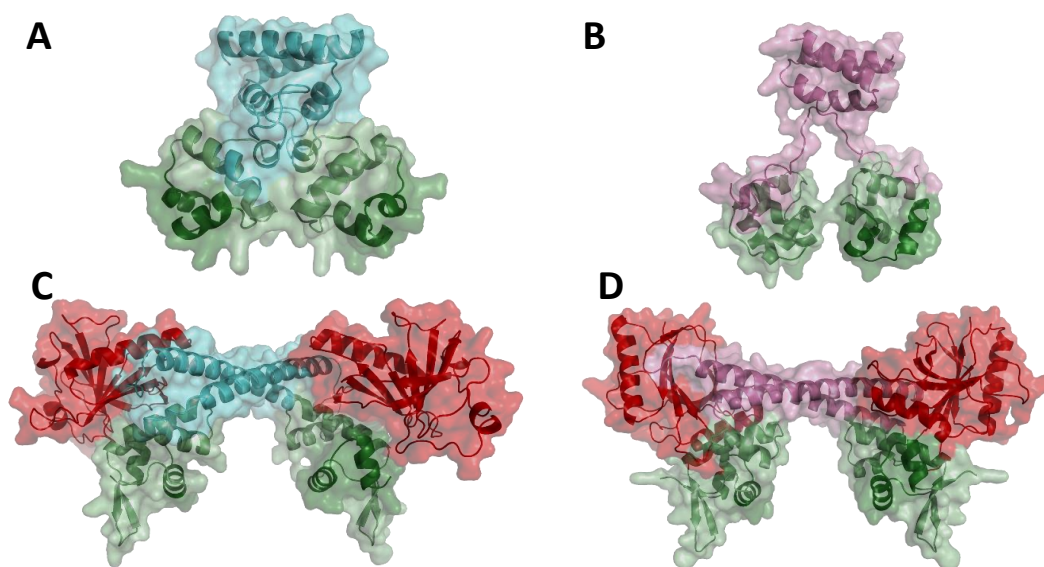


**Figura 8. Predicción de la estructura terciaria de las proteínas Alr1976 (A) y All0345 (B) mediante SwissModel.** En ambos casos los modelos construidos por SwissModel están superimpuestos sobre la estructura predicha por AlphaFold (color crema), en la que se indica en color verde el dominio de unión a DNA, con el motivo HTH resaltado en verde oscuro. En el caso de la proteína All0345 se indica además en rojo el dominio de unión a efector de tipo AraC-E. **A.** Modelado por homología de la estructura de Alr1976. Se representa en color azul la estructura modelada en base al regulador BswR de *Pseudomonas aeruginosa* (PDB: 4O8B) y en rosa la modelada en base al regulador EspR de *Mycobacterium tuberculosis* (PDB: 3QF3). **B.** Modelado por homología de la estructura de All0345. Se representa en color azul la estructura modelada en base al regulador EcmrR de *Escherichia coli* (PDB: 6XLK) y en rosa la modelada en base al regulador BmrR de *Bacillus subtilis* (PDB: 1R8E).

Se construyeron dos modelos de Alr1976. En uno se empleó como molde la estructura del regulador BswR de *Pseudomonas aeruginosa* (PDB: 4O8B) (69), un regulador transcripcional que responde a la presencia de xenobióticos y juega un papel muy importante en la movilidad y la formación de biofilms, y que tenía un 18,81 % de identidad y un 83 % de cobertura. En otro se empleó la estructura del regulador EspR de *Mycobacterium tuberculosis* (PDB: 3QF3) (73), que presentaba un 15,38 % de identidad y un 86 % de cobertura (Figura 8A) y está relacionado con la virulencia de este patógeno (74, 75). En ambos casos, la estructura del dominio de unión a DNA coincide con la predicción de AlphaFold, pues se predice que se compone de 5 hélices  $\alpha$  separadas por bucles que tienen la misma orientación. No obstante, a diferencia de la predicción de AlphaFold, el modelado por homología no predice la existencia de láminas  $\beta$  al principio y al final del dominio de unión a DNA. Además, los dos modelos por homología, al igual que AlphaFold, predicen la existencia de dos hélices  $\alpha$  en el extremo C-terminal, aunque su orientación espacial varía.

Con respecto a la proteína All0345 se construyó un modelo con la estructura del regulador EcmrR de *Escherichia coli* (PDB: 6XLK) (76), con un 29,77 % de identidad y un 90 % de cobertura, y otro con la estructura del regulador BmrR de *Bacillus subtilis* (PDB: 1R8E) (72), con un 24,48 % de identidad y un 93 % de cobertura (Figura 8B), que en ambos casos son activadores transcripcionales implicados en la resistencia diversos fármacos (72, 76). Se observa que los dos modelos predicen la misma estructura que AlphaFold para el dominio de unión a DNA, con cuatro hélices  $\alpha$  y una lámina  $\beta$  que muestran la misma orientación espacial, así como una larga hélice  $\alpha$  que separa ambos dominios. Además, aunque la orientación espacial del dominio de unión a efector es diferente, en ambos modelos se predice que tiene la misma estructura tridimensional, compuesta de dos hélices  $\alpha$  y una hoja  $\beta$  antiparalela formada por ocho láminas  $\beta$ .

Además, el modelado mediante SwissModel sugería que los reguladores Alr1976 y All0345 eran capaces de dimerizar, lo que permitió predecir estructuras de dímeros para ambos (Figura 9).



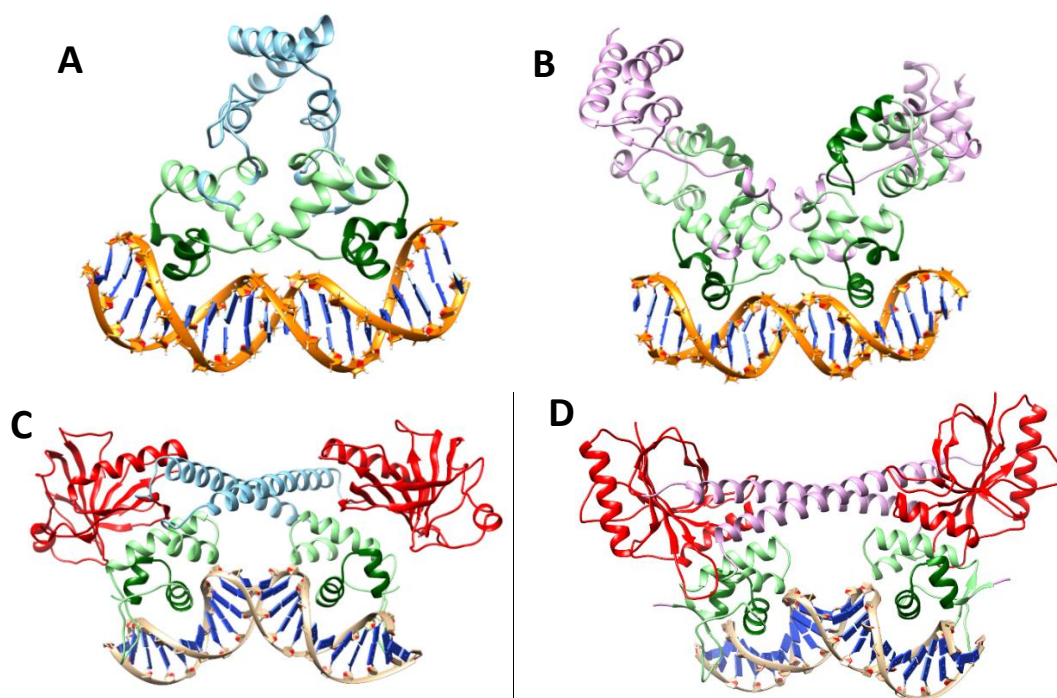
**Figura 9. Predicción de dímeros de las proteínas Alr1976 (A, B) y All0345 (C, D) mediante el servidor SwissModel.** En ambos casos se mostró la superficie con una transparencia del 50 % para visualizar los motivos de estructura secundaria. Se indica en color verde el dominio de unión a DNA, con el motivo HTH resaltado en verde oscuro, y en el caso de la proteína All0345 se indica además en rojo el dominio de unión a efector de tipo AraC-E. Se modelaron los dímeros de Alr1976 en base al regulador BswR de *P. aeruginosa* (PDB: 4O8B) (A) y en base al regulador EspR de *M. tuberculosis* (PDB: 3QF3) (B) y los dímeros de All0345 en base al regulador EcmrR de *E. coli* (PDB: 6XLK) (C) y en base al regulador BmrR de *B. subtilis* (PDB: 1R8E) (D)



En la figura 9, se puede la estructura predicha para los dímeros de Alr1976 y All0345. En el caso del regulador Alr1976 se observa que los dímeros modelados en base al regulador BswR de *P. aeruginosa* forman una estructura compacta en la que los dominios de unión a DNA quedan muy próximos (figura 9A) mientras que los dímeros modelados en base al regulador EspR de *M. tuberculosis* son menos compactos y los dominios de unión a DNA están más separados (figura 9B). En ambas predicciones, las hélices  $\alpha$  del extremo carboxilo terminal están entrelazadas, y las caras internas de las hélices  $\alpha$  terminales son ricas en residuos hidrofóbicos, que podrían mediar la dimerización (Material Suplementario, sección, 9.6, figura S-1).

Con respecto al regulador All0345 se observa que la predicción de dímeros en base al regulador EcmrR de *E. coli* y al regulador BmrR de *B. subtilis* son bastante similares (figura 9 C y D). Tanto los dominios de unión a DNA como los dominios de unión a efector quedan expuestos, lo que les permitiría interaccionar con el DNA y con los posibles efectores, respectivamente, y que en este caso los dominios de unión a DNA están más separados entre sí que en los dímeros del regulador Alr1976. En ambos modelos, el regulador All0345 dimerizaría a través de la larga hélice  $\alpha$  que conecta el dominio de unión a DNA con el dominio efector, la cual es rica en residuos hidrofóbicos en su cara interna, los cuales podrían mediar la dimerización a través de interacciones hidrofóbicas (Material Suplementario, sección, 9.6, figura S-2).

Dado que se han resuelto o modelado los complejos proteína-DNA de las estructuras empleadas para modelar los dímeros de ambos reguladores (69, 72, 73, 76), se decidió alinear los modelos de dímeros sobre estos complejos para modelar complejos proteína-DNA (Figura 10)



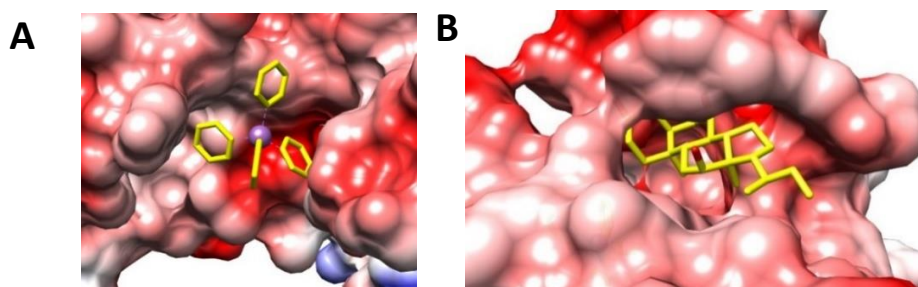
**Figura 10. Predicción de la unión al DNA de las proteínas Alr1976 (A, B) y All0345 (C, D) realizadas en Chimera.** Se indica en color verde el dominio de unión a DNA, con el motivo HTH resaltado en verde oscuro, y en el caso de la proteína All0345 se indica además en rojo el dominio de unión a efector de tipo AraC-E. Se modelaron los complejos de Alr1976 con el DNA en base al regulador BswR de *P. aeruginosa* (PDB: 4O8B) (A) y en base al regulador EspR de *M. tuberculosis* (PDB: 3QF3) (B) y los complejos de All0345 y DNA en base al regulador EcmrR de *E. coli* (PDB: 6XLK) (C) y en base al regulador BmrR de *B. subtilis* (PDB: 1R8E) (D).

Como se observa en la figura 10, los modelos DNA-proteína no presentan choques estéricos entre la proteína y el DNA. En el caso de Alr1976, cuando el complejo proteína-DNA se modela en base a BswR de *P. aeruginosa* (figura 10A), se aprecia como la interacción del regulador con el DNA tiene lugar principalmente a través del dominio HTH de cada dímero y la hélice  $\alpha$  número 3, se encuentra insertada en el surco mayor del DNA. Por el contrario, cuando se modela en base a EspR de *M. tuberculosis* (figura 10B) se aprecia que la unidad funcional que se une al DNA es un dímero de dímeros. En este modelo solo interacciona con el DNA un motivo HTH de cada dímero, quedando por tanto dos motivos HTH de unión a DNA libres, que en el caso del regulador EspR se ha propuesto que podrían interaccionar con regiones del DNA más distantes (73).

En el caso de Alr0345, tanto en el modelo construido en base a EcmrR de *E. coli* (figura 10C) como en el construido en base a BmrR de *B. subtilis* (figura 10D), se aprecia como la interacción con el DNA tiene lugar a través de los motivos HTH del dominio de unión a DNA, cuyas hélices se introducen en el surco mayor del DNA. Además, se observa que en estos modelos la doble hélice del DNA está deformada, lo que se ha descrito que permite que los elementos -35 y -10 de la región promotora queden próximos y puedan ser reconocidos por el factor  $\sigma$  de la RNA polimerasa para iniciar la transcripción (76).

En todos los modelos, tanto de Alr1976 como de Alr0345, las cargas positivas se concentran en los dominios de unión al DNA, para facilitar su interacción con este a través de sus grupos fosfatos, mientras que las cargas negativas y neutras se encuentran distribuidas a lo largo del resto de la proteína. (Material Suplementario, sección 9.7)

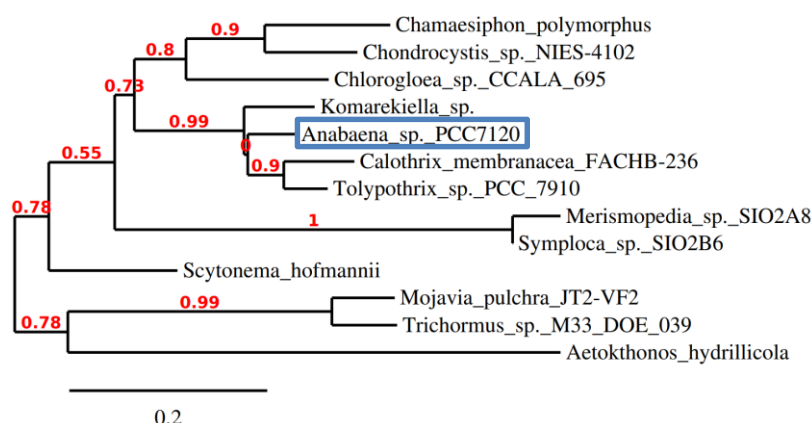
Finalmente, en la estructura del regulador EcmR de *E. coli* se ha descrito la existencia de dos sitios de unión para ligandos, que están ocupados por dos moléculas usadas en la preparación de la muestra, el tetrafenilantiamonio (TPSb<sup>+</sup>) y 3- ([3-colamidopropilo]dimetilamonio)-2-hidroxi-1-propanosulfonato (CHAPSO) (76). Por tanto, se decidió superponer el modelo del regulador Alr0345 con la estructura del regulador EcmR de *E. coli* y estudiar si los ligandos del regulador EcmR podían alojarse en la estructura del Alr0345 (figura 11).



**Figura 11. Modelo de la interacción de Alr0345 con TPSb<sup>+</sup> (A) y con CHAPSO (B).** El modelo del regulador Alr0345 en base al regulador EcmR se superpuso con la estructura de EcmR en complejo con ligandos (PDB: 6XLK) y se mostraron las estructuras del TPSb<sup>+</sup> y CHAPSO, que se resaltaron en color amarillo

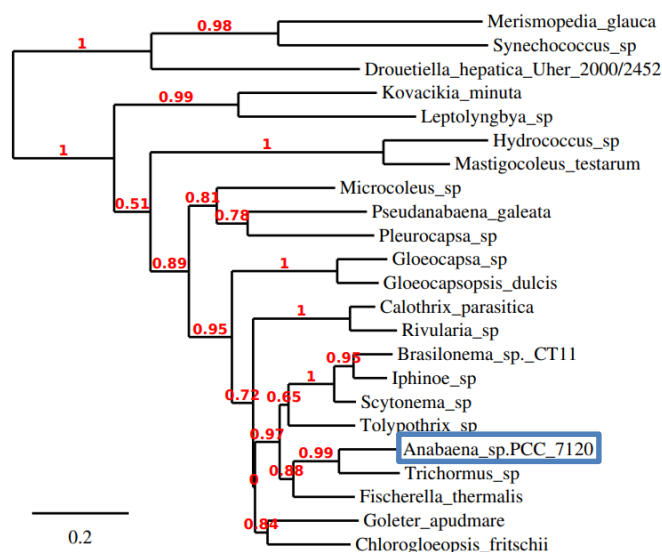
Como se aprecia en la figura 10, el modelo del regulador Alr0345 presenta dos bolsillos en el dominio C-terminal que permitirían la interacción con ligandos, pues tanto el TPSb<sup>+</sup> (Figura 11A) como el CHAPSO (Figura 11B) que aparecen en la estructura del regulador EcmR pueden alojarse en el modelo del regulador Alr0345 sin impedimentos estéricos.

A continuación, para conocer el grado de conservación de las proteínas de interés en cianobacterias se buscaron homólogos en otras cianobacterias y se construyeron árboles filogenéticos que se muestran las figuras 12 y 13.



**Figura 12.** Árbol filogenético de la proteína Alr1976 con homólogos en otras cianobacterias. Se recuadra en azul la secuencia de Alr1976 de *Anabaena* sp. PCC7120. Los números en rojo son el resultado de la validación por bootstrap.

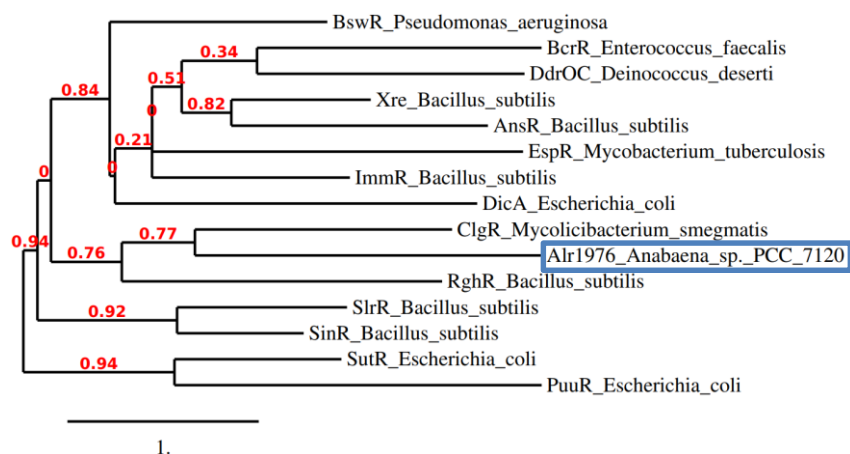
En la figura 12, se puede apreciar como la proteína Alr1976 es filogenéticamente más próxima a las secuencias de *Komarekiella* y *Calothrix*, filamentosas y formadoras de heterocistos, aunque también hay homólogos en cianobacterias no filamentosas como *Chamaesiphon*. El alineamiento múltiple de estas secuencias mostró que mientras el dominio de unión a DNA de tipo XRE está muy conservado en todas las secuencias analizadas, el dominio C-terminal está menos conservado (Material Suplementario, sección 9.9, figura S-5).



**Figura 13.** Árbol filogenético de la proteína All0345 con homólogos en otras cianobacterias. Se recuadra en azul la secuencia de All0345 de *Anabaena* sp. PCC7120. Los números en rojo son el resultado de la validación por bootstrap.

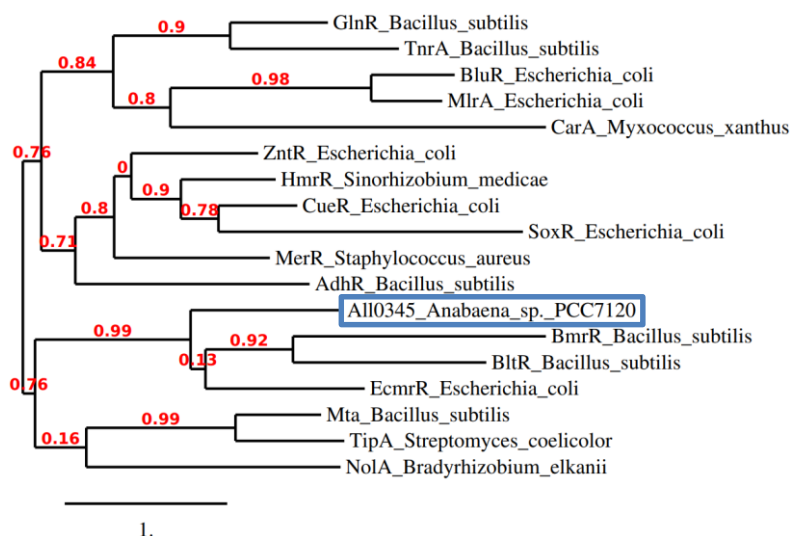
Con respecto al regulador All0345, se puede observar que el árbol filogenético de las secuencias de cianobacterias (figura 13), está formado por dos clados. En el clado superior se encuentran las secuencias de *Merismopedia glauca*, *Synechococcus* y *Drouetiella hepática*, todas ellas no fijadoras de nitrógeno, mientras que en el clado inferior se encuentran el resto de secuencias analizadas, incluyendo la de *Anabaena*, que se encuentra filogenéticamente más próxima a las secuencias *Trichormus* sp. y *Fischerella thermalis*, ambas cianobacterias filamentosas y diazotróficas. El alineamiento múltiple de estas secuencias puso de manifiesto que la proteína All0345 se encuentra altamente conservada en cianobacterias, tanto fijadoras como no fijadoras, teniendo conservado ambos dominios, tanto el de unión a DNA tipo MerR, como el de unión a efector de tipo AraC-E (Material Suplementario, sección 9.9, figura S-6).

Con el objetivo de obtener información con respecto a la función desarrollada por las proteínas Alr1976 y All0345, se estudió su similitud con proteínas ya caracterizadas de la misma familia en otros procariotas. Para ello, en el caso de Alr1976, se buscaron secuencias de reguladores de la familia XRE, mientras que para All0345, de la familia MerR y se construyó un árbol filogenético (figuras 14 y 15).



**Figura 14.** Árbol filogenético de la proteína Alr1976 comparando secuencias de proteínas no cianobacterianas obtenidas tras realizar una búsqueda en Uniprot de proteínas pertenecientes a la familia XRE. Se recuadra en azul la secuencia de Alr1976 de *Anabaena* sp. PCC7120. Los números en rojo son el resultado de la validación por bootstrap.

En la figura 14, se aprecia como las proteínas con secuencia más similares a Alr1976 son RghR de *Bacillus Subtilis*, el cual está implicado en la regulación del operón *srfA*, involucrado en la producción de la surfactina y en la esporulación (77) y ClgR de *Mycobacterium smegmatis*, que participa en la respuesta a estreses. El alineamiento múltiple mostró que la conservación de la proteína Alr1976 en procariotas es mucho menor que en cianobacterias (Material suplementario, sección 9.11 figura S-7). Solo el dominio de unión a DNA de tipo XRE está conservado y la similitud entre secuencias es mucho menor. Por el contrario, en los filogenéticamente más cercanos a Alr1976 (RghR de *B. subtilis* y ClgR de *M. smegmatis*), toda la secuencia proteica está altamente conservada (Material Suplementario, sección 9.11, figura S-8)



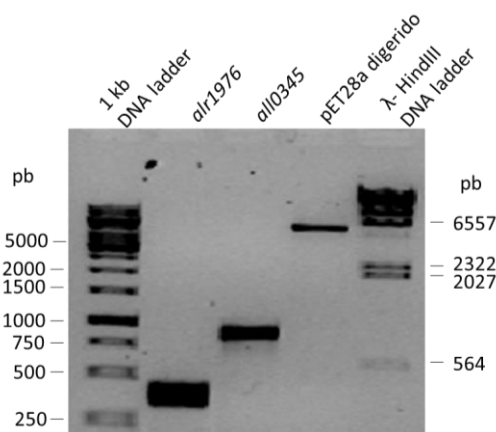
**Figura 15.** Árbol filogenético de la proteína All0345 comparando secuencias de proteínas no cianobacterianas obtenidas tras realizar una búsqueda en Uniprot de proteínas pertenecientes a la familia MerR. Se recuadra en azul la secuencia de All0345 de *Anabaena* sp. PCC7120. Los números en rojo son el resultado de la validación por bootstrap.



Como se muestra en la figura 15, en el caso del regulador All0345, las proteínas filogenéticamente más cercanas son BmrR de *Bacillus subtilis* que es un regulador del transportador de fármacos Bmr (72) y EcmrR de *E. coli*, cuya unión al DNA está modulada por fármacos, aunque su función es desconocida (76). En procariotas, se tiene una conservación únicamente del dominio de unión al DNA de tipo MerR, mientras que el dominio de unión a efector de tipo AraC-E no está conservado (Material Suplementario, Sección 9.11, figura S-9). Por el contrario, en aquellos filogenéticamente cercanos (EcmrR de *E. coli* y BltR y BmrR de *B. subtilis*) sí que están conservados tanto el dominio de unión a DNA de tipo MerR como el dominio de unión a efector de tipo AraC-E (Material Suplementario, Sección 9.11, figura S-10)

## 5.2. Clonaje en pET-28a

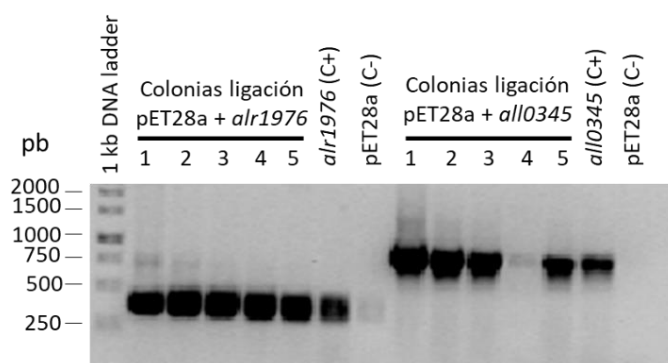
Tras analizar sus propiedades mediante estudios bioinformáticos se decidió purificar las proteínas de interés mediante sobreexpresión heteróloga en células de *E. coli*, para lo cual se clonaron los genes *alr1976* y *all0345* en el plásmido pET28a con una cola de histidinas en el extremo amino terminal. Para realizar el clonaje se amplificaron las secuencias codificantes de los genes y se introdujeron sitios de restricción mediante los oligonucleótidos utilizados para la amplificación. En concreto, se introdujo el sitio NdeI en su extremo 5' y el sitio SalI en su extremo 3' para así poder clonar estos genes en el sitio de clonaje múltiple del pET28a de forma que no se alterase el marco de lectura y se incorporase la cola de histidinas en el extremo amino terminal. Tras amplificar los genes, se llevó a cabo una doble digestión de los insertos y del plásmido con las enzimas SalI y NdeI y para verificar que tanto la amplificación de los insertos como la digestión vector habían sido correctas, se realizó una electroforesis en un gel de agarosa, cuyo resultado se muestra en la figura 16.



**Figura 16. Análisis de los insertos y el vector utilizados para el clonaje.** Electroforesis en gel de agarosa al 1 % de los genes *alr1976* y *all0345* amplificados por PCR y digeridos con las enzimas de restricción SalI y NdeI (carriles 2 y 3, respectivamente) y el vector pET28a digerido con las enzimas de restricción SalI y NdeI (carril 4). El carril 1 contiene el marcador de pesos moleculares 1 kb DNA ladder y el 5, el marcador  $\lambda$ -HindIII.

Como se aprecia en la figura 16, ambos insertos amplificaron correctamente, ya que se obtiene una banda entre los 200 y 500 pares de bases para el gen *alr1976*, el cual tiene un tamaño de 366 pares de bases y una banda entre 750 y 1000 para el gen *all0345*, que tiene una longitud de 873 pares de bases. Por otro lado, se puede observar la correcta digestión del plásmido pET28a, pues se observa una banda entre 5 y 6 kb y el tamaño del plásmido es de 5,4 kb.

Tras verificar la correcta amplificación y digestión de los insertos y el vector, se llevó a cabo una ligación de los insertos *alr1976* y *all0345* en el vector pET28a haciendo uso de los extremos cohesivos generados por las enzimas de restricción. Tras la ligación se llevó a cabo un choque térmico de las mezclas de ligación pET28a + *alr1976* y pET28a + *all0345* en células DH5 $\alpha$ . Para determinar qué colonias habían incorporado el vector con el inserto de interés se llevó a cabo una PCR colonia empleando parejas de cebadores que amplificaban la secuencia codificante de los genes *alr1976* y *all0345*. Se analizaron 5 colonias obtenidas en el choque térmico de la ligación pET28a + *alr1976* y 5 colonias de la ligación pET28a + *all0345* en un gel de agarosa al 1%, empleando como controles positivos los genes *alr1976* y *all0345* y como control negativo el vector pET28a vacío. Los resultados de la electroforesis se muestran en la figura 17.

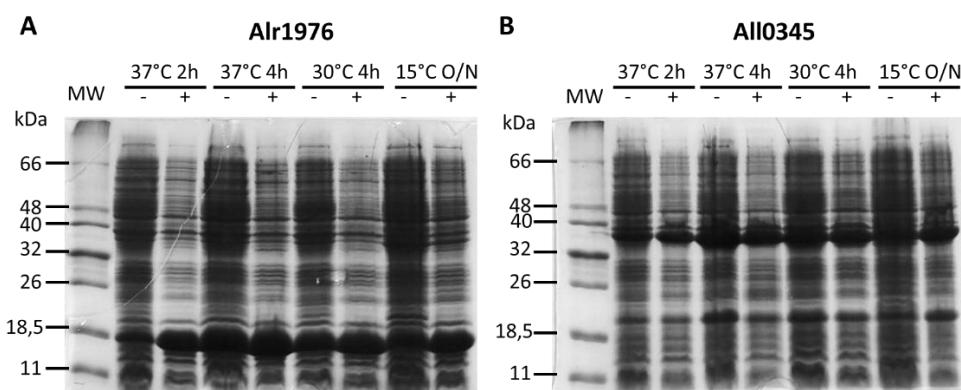


**Figura 17. PCR colonia de la ligación de los insertos *alr1976* y *all0345* en el vector pET28a.** Análisis mediante electroforesis en gel de agarosa al 1 % de los productos de amplificación obtenidos tras realizar una PCR colonia para comprobar la presencia del inserto de interés en las colonias de *E. coli* DH5 $\alpha$  obtenidas en el choque térmico de las mezclas de ligación pET28a + *alr1976* y pET28a + *all0345*. El primer carril corresponde al marcador de pesos moleculares 1 kb DNA ladder, los carriles 2-6, a las colonias obtenidas tras el choque térmico con la mezcla de ligación pET28a + *alr1976* y los carriles 9-13 a las colonias obtenidas tras el choque térmico con la mezcla de ligación pET28a + *all0345*. Los carriles 7 y 14 corresponden a la amplificación sobre los genes *alr1976* y *all0345* respectivamente (controles positivos) y los carriles 8 y 15 a la amplificación sobre el vector pET28a sin inserto (control negativo).

Como se puede observar en la figura 17, todas las colonias transformadas con la mezcla de ligación pET28a + *alr1976* y 4 de las 5 colonias transformadas con la ligación pET28a + *all0345* analizadas contenían el vector con el inserto, pues se observa que no se obtiene producto de amplificación al llevar a cabo la reacción de amplificación sobre el vector pET28a vacío. En ambos casos, se seleccionó la colonia 1 y se iniciaron preinóculos en medio LB con kanamicina, que se incubaron toda la noche a 37 °C. Al día siguiente se aislaron los plásmidos mediante miniprep y se secuenciaron para comprobar que los insertos no contenían mutaciones, obteniendo así los plásmidos pET28a:*alr1976* y pET28a:*all0345*. Finalmente, los plásmidos se introdujeron en células BL21 termocompetentes mediante choque térmico, para generar las estirpes BL21 pET*alr1976* y BL21 pET*all0345*, que se usaron para la sobreexpresión de las proteínas.

### 5.3. Prueba de sobreexpresión

Para poder purificar las proteínas Alr1976 y All0345 se optó por su sobreexpresión en las cepas BL21 pETalr1976 y BL21 pETall0345. Para conocer las condiciones de máxima expresión de las proteínas Alr1976 y All0345, se realizó una prueba de sobreexpresión a diferentes tiempos y temperaturas de inducción: 2 h a 37 °C, 4 h a 37 °C, 4 h a 30 °C y toda la noche a 15 °C, manteniendo constante la concentración de inductor, en este caso 1 mM de IPTG. Una vez transcurrido el tiempo de inducción, se recogieron las muestras de cada cultivo y se analizaron mediante electroforesis desnaturalizante en geles de SDS-PAGE, cuyos resultados se recogen en la figura 18.



**Figura 18. Prueba de sobreexpresión de las proteínas Alr1976 (A) y All0345 (B) en *E. coli* BL21** En ambos geles, el primer carril (MW) corresponde al marcador de pesos moleculares Low Molecular Weight Protein Marker II (NZY) y el resto de los carriles a los extractos celulares sin inducir (-) e inducidos (+) obtenidos en las diferentes condiciones de inducción estudiadas (2 h a 37 °C, 4 h a 37 °C, 4 h a 30 °C y toda la noche a 15 °C). El análisis se llevó a cabo mediante electroforesis desnaturalizante en gel de poliacrilamida al 17 %.

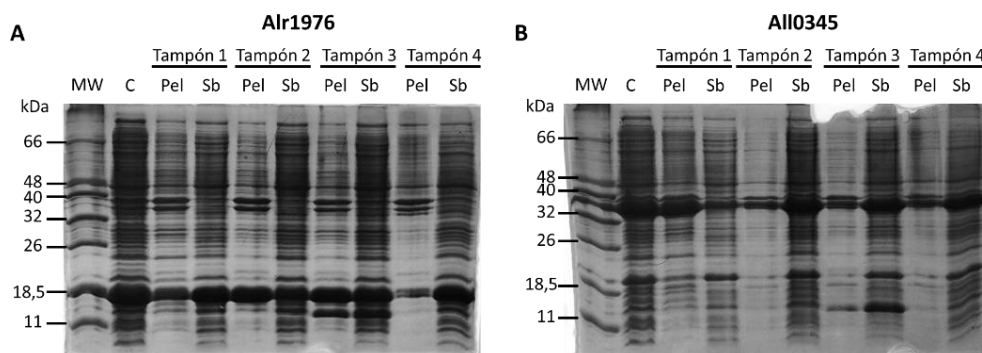
Como se observa en la figura 18, en todas las condiciones analizadas hay una sobreexpresión considerable de las proteínas Alr1976 y All0345. Con los resultados obtenidos en la prueba de sobreexpresión, se decidió llevar a cabo la sobreexpresión de las proteínas Alr1976 y All0345 a gran escala incubando a 15 °C durante toda la noche tras la adición del IPTG para facilitar el plegamiento, reduciendo así la formación de posibles cuerpos de inclusión.

### 5.4. Pruebas de solubilización

Una vez obtenida la biomasa de células BL21 pETalr1976 y BL21pETall0345 sobreexpresando las proteínas de interés, es necesario extraer las proteínas del interior celular y solubilizarlas, para lo cual se decidió lisar las células mediante sonicación. Antes de sonicar la biomasa obtenida se realizó una prueba de solubilidad a pequeña escala de las proteínas en diferentes tampones, para así encontrar el tampón en el que solubilizara la mayor parte de la proteína de interés y que estas no precipitasen. Se estudiaron 4 tampones diferentes: el primero estaba compuesto por Tris 50 mM pH 8, 10 % de glicerol y 500 mM de NaCl; el segundo por Tris 50 mM pH 8, NaCl 200 mM, 10 % de glicerol y Tween 0'05 %; el tercero por Tris 50 mM, pH 8 y 1 U de lisozima; y el último por NaH<sub>2</sub>PO<sub>4</sub> 0,1 M, Tris 0,01 M y Cloruro de guanidinio 2 M a pH 8.

Para llevar a cabo esta prueba se descongeló una pequeña fracción de la biomasa, la cual fue sometida a lisis celular en los tampones anteriores y posteriormente se centrifugó para separar la fase soluble (sobrenadante) de la fase insoluble (precipitado) y determinar si la proteína de interés era soluble mediante electroforesis desnaturalizante en gel de poliacrilamida SDS-PAGE, cuyos resultados se muestran en la figura 19.





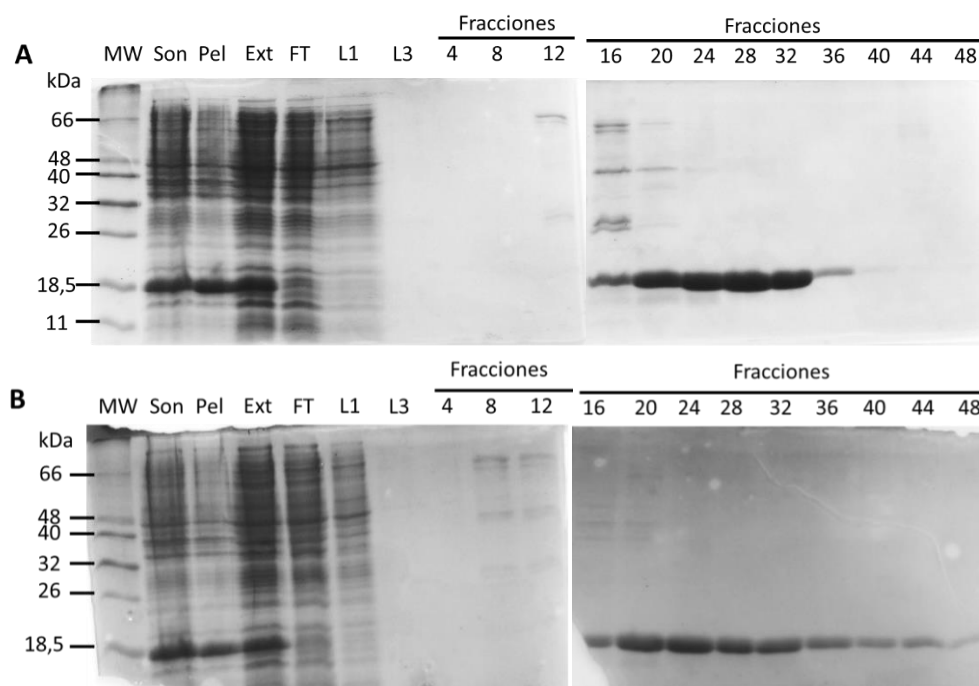
**Figura 19. Resultados de la prueba de solubilidad de las proteínas Alr1976 (A) y All0345 (B) en diferentes tampones analizados.** En ambos geles, el primer carril (MW) corresponde al marcador de pesos moleculares Low Molecular Weight Protein Marker II (NZY), el segundo carril (C) corresponde a un extracto de proteínas totales de la célula mediante lisis con SDS y los carriles restantes son el precipitado (Pel) y sobrenadante (Sb) obtenidos tras la lisis celular por sonicación de la biomasa de células BL21-pET $_{Alr1976}$  y pET $_{All0345}$  resuspendida en los diferentes tampones: tampón 1 (Tris 50 mM pH 8, 10 % de glicerol y 500 mM de NaCl), tampón 2 (Tris 50 mM pH 8, NaCl 200 mM, 10 % de glicerol y Tween 0'05 %), tampón 3 (Tris 50 mM, pH 8 y 1 U de lisozima) y tampón 4 ( $\text{NaH}_2\text{PO}_4$  0,1 M, Tris 0,01 M, Cloruro de guanidinio 2 M, pH 8). El análisis se llevó a cabo mediante electroforesis desnaturizante en gel de poliacrilamida al 17 %.

Como se observa en la figura 19, la solubilidad de las proteínas Alr1976 y All0345 es diferente según el tampón empleado. Para el regulador Alr1976 se observa que la solubilidad es mucho mayor en el en el tampón 4, el cual contiene cloruro de guanidinio, un agente caotrópico que a la concentración de trabajo (2 M) se asocia a las cadenas laterales de los aminoácidos, minimizando las interacciones que dan lugar a los fenómenos de agregación (78). No obstante, se observa que la proteína también se solubiliza en el resto de condiciones estudiadas. Respecto al regulador All0345, se puede apreciar como su solubilidad en el tampón 1 es baja mientras que en el resto de tampones (2, 3 y 4), la solubilidad de All0345 es más alta. A la vista de los resultados obtenidos, en ambos casos se eligió el tampón 2 (Tris 50 mM pH 8, NaCl 200 mM, 10 % de glicerol y Tween 0'05 %) para llevar a cabo la purificación de las proteínas Alr1976 y All0345, pues en ambos casos se producía una buena solubilización de las proteínas de interés.

### 5.5. Purificación mediante cromatografía de afinidad a iones metálicos inmovilizados

La purificación de ambos reguladores se llevó a cabo mediante cromatografía de afinidad a iones metálicos inmovilizados (IMAC), ya que las proteínas recombinantes contenían una cola de histidinas en su extremo amino terminal que muestra afinidad por estos iones.

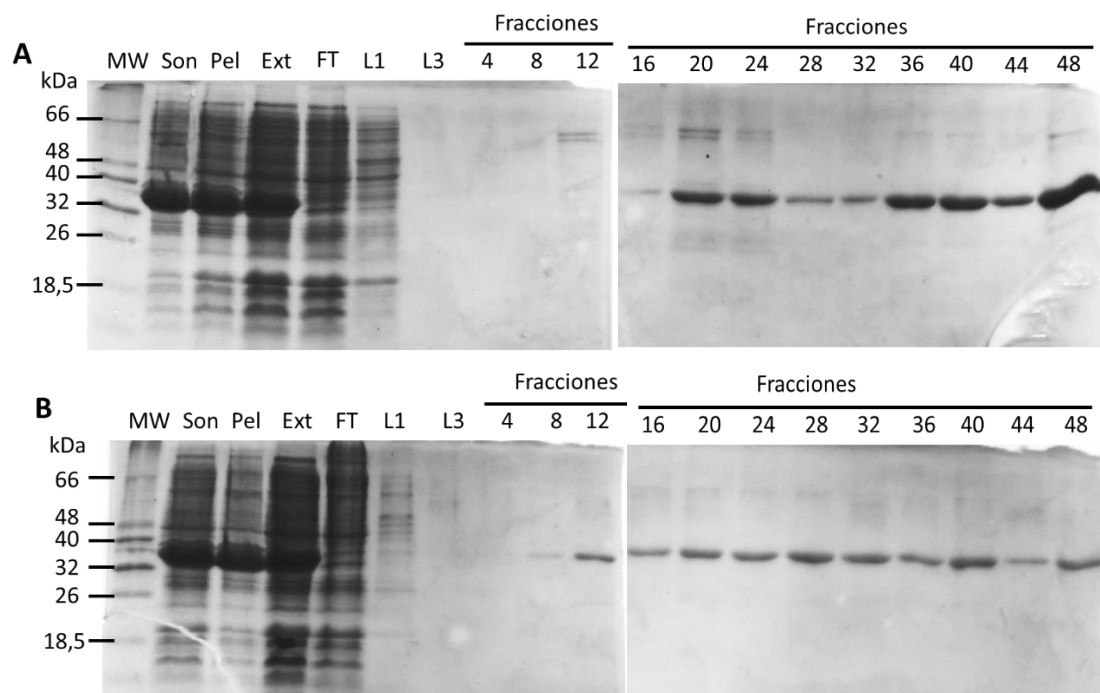
Las proteínas Alr1976 y All0345 se sonicaron en tampón compuesto por Tris 50 mM, Tween 0,05 %, glicerol 10 %, NaCl 200 mM e imidazol 20 mM a pH 8, y se llevaron a cabo dos purificaciones, una empleando cobalto y otra empleando níquel como metales fijados en la columna, pues la elución de las proteínas fijadas a la columna es diferente según el metal empleado. Para evaluar la eficacia de cada etapa de la purificación se recogieron alícuotas tras la lisis de las células mediante sonicación, del precipitado obtenido tras clarificar el sonicado (*pellet*), del extracto crudo antes de cargarlo en la columna, del volumen eluido al cargar el extracto crudo en la columna (*flowthrough*) y de los lavados. Estas alícuotas junto a las fracciones obtenidas en la elución de la columna mediante el gradiente de imidazol se precipitaron con TCA 10 % y se analizaron mediante electroforesis desnaturizante en gel de poliacrilamida SDS-PAGE, cuyos resultados se muestran en las figuras 20 y 21.



**Figura 20. Perfil de la elución de la columna de purificación de la proteína Alr1976, empleando níquel (A) y cobalto (B) como ion metálico inmovilizado.** En ambos geles, el primer carril (MW) corresponde al marcador de pesos moleculares Low Molecular Weight Protein Marker II (NZY), el segundo carril (Son) corresponde a una alícuota recogida tras la sonicación, el tercero (Pel) al precipitado obtenido tras clarificar el sonicado, el cuarto (Ext) al extracto crudo, el quinto (FT) al volumen eluido al cargar el extracto crudo, el sexto (L1) al primer lavado, el séptimo (L3) al tercer lavado y los carriles restantes (F4-F48) a las fracciones obtenidas tras la elución de la columna con un gradiente de imidazol (0-1 M). El análisis se llevó a cabo mediante electroforesis desnaturizante en gel de poliacrilamida al 17 %.

Como se aprecia en la figura 20, en los carriles 2 y 4 (sonicado y extracto crudo) aparece una banda en torno a 18,5 KDa correspondiente a la proteína Alr1976, que es indicativa de una buena lisis celular y solubilización de la proteína, aunque parte de ella queda en el precipitado obtenido tras clarificar (carril 3). En el carril correspondiente al *flowthrough*, se aprecia como dicha banda no aparece, lo que indica que la proteína se ha fijado a la columna, y aparecen otras muchas bandas, correspondientes a otras proteínas de la célula. En el carril del primer lavado, se sigue apreciando la presencia de otras proteínas, pero no así de la proteína Alr1976 y en el carril correspondiente al tercer lavado, apenas aparecen proteínas, lo que sugiere que ya se han eluido todas aquellas proteínas que no presentan afinidad por el ion metálico. En el caso de la purificación con níquel (figura 20A), la proteína Alr1976 aparece a partir de la fracción 16, en la cual además se aprecia la presencia de otras proteínas contaminantes. Las fracciones 20-32 presentan una alta concentración de proteína y ausencia de contaminantes y la última fracción de aquellas analizadas que contiene la proteína de interés es la 36, en la que la concentración de Alr1976 es menor que en las fracciones anteriores.

En el caso de purificación con cobalto (figura 20B), el patrón de elución es diferente. En este caso, la proteína Alr1976 aparece a partir de la fracción 16 y hasta la fracción 48. No obstante, en las fracciones 16-32, la intensidad de la banda a 18,5 KDa es mayor que en las fracciones 36-48, lo que indica que en estas primeras fracciones la concentración de Alr1976 es mucho mayor. Finalmente, se puede apreciar que en la purificación con cobalto las fracciones presentan menos contaminantes que en la purificación con níquel.



**Figura 21. Perfil de la elución de la columna de purificación de la proteína All0345, empleando níquel (A) y cobalto (B) como ion metálico inmovilizado.** En ambos geles, el primer carril (MW) corresponde al marcador de pesos moleculares Low Molecular Weight Protein Marker II (NZY), el segundo carril (Son) corresponde a una alícuota recogida tras la sonicación, el tercero (Pel) al precipitado obtenido tras clarificar el sonicado, el cuarto (Ext) al extracto crudo, el quinto (FT) al volumen eluido al cargar el extracto crudo, el sexto (L1) al primer lavado, el séptimo (L3) al tercer lavado y los carriles restantes (F4-F48) a las fracciones obtenidas tras la elución de la columna con un gradiente de imidazol (0-1 M). El análisis se llevó a cabo mediante electroforesis desnaturalizante en gel de poliacrilamida al 17 %.

La figura 21 muestra el resultado de la purificación del regulador All0345. Se observa que, al igual que sucedía para el regulador Alr1976, los carriles correspondientes al sonicado (Son) y al extracto (Ext) presentan una gruesa banda entre 32 y 40 KDa, lo que indica que, a pesar de que una parte de la proteína se encuentra en el precipitado obtenido tras la sonicación (Pel), se ha producido la liberación de la proteína de la célula y permanece soluble en el extracto crudo. Además, al cargar el extracto en la columna de IMAC, la proteína se fija a la columna, tanto cuando se usa el cobalto como el níquel como ion metálico, ya que la banda correspondiente a All0345 no aparece en el carril del volumen eluido al cargar el extracto crudo (FT). Los carriles de los lavados, muestran que se logra eluir las proteínas que unidas por interacciones débiles a la columna sin que eluya la proteína de interés.

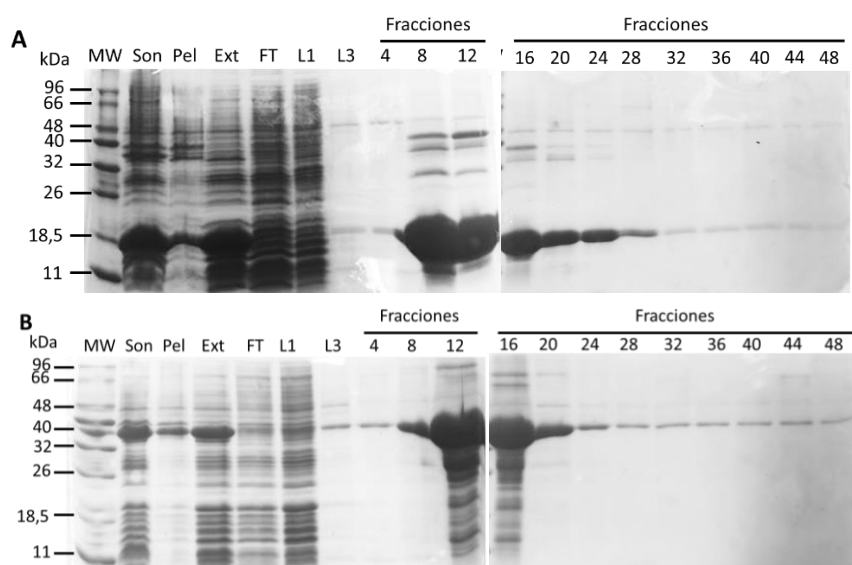
En la purificación empleando como ion metálico el níquel (figura 21A), la proteína All0345 comienza a eluir a partir de la fracción número 20, aunque no es hasta la fracción 36 cuando su concentración aumenta notablemente. Además, en las primeras fracciones en las que aparece la proteína (20-24) se observa la presencia de contaminantes, ya que se aprecian bandas tanto de pesos moleculares mayores como de pesos menores que la proteína de interés en el gel de electroforesis. Respecto a la purificación llevada a cabo con cobalto (figura 21B), la proteína All0345 eluye desde la fracción número 12 hasta la fracción 48, en concentraciones similares en todas las fracciones y con menor presencia de otras proteínas contaminantes que en la purificación con níquel.

Una vez analizadas las fracciones, se seleccionaron aquellas que contenían las proteínas de interés en mayor cantidad y con mayor grado de pureza y se sometieron a diálisis para poder eliminar el imidazol de las alícuotas de proteína. Para ello, se combinaron las fracciones seleccionadas y se sometieron a dos ciclos de diálisis a 4 °C en 3 L de tampón Tris 50 mM, glicerol 10 % (v/v) y NaCl 100 mM, a pH 8. Para la proteína Alr1976, de la purificación con níquel, se combinaron las fracciones 26 a 36 y 22 a 25 y para la purificación con cobalto se seleccionaron las fracciones 16 a 22 y 23 a 48. Por otro lado, para la proteína All0345 en el caso de la purificación con níquel se combinaron las fracciones de la 36 a la 48 y en el caso de la purificación con cobalto, las fracciones 12 a 48.

Tras los ciclos de diálisis las proteínas precipitaron en la membrana de diálisis. Por ello se centrifugó el contenido para eliminar las proteínas precipitadas y obtener la proteína que permaneciera en solución. Finalmente, se llevó a cabo un espectro UV-visible y a partir de la absorbancia a 280 nm y el coeficiente de extinción teórico, se pudo calcular la concentración a la que se encontraban las proteínas, siendo esta muy baja (en torno a 5 µM) debido a la precipitación que tuvo lugar durante la diálisis.

Dado que mediante la purificación empleando el tampón Tris 50 mM no se consiguió purificar la proteína en cantidad suficiente para poder llevar a cabo su caracterización se decidió llevar a cabo una purificación empleando un tampón con cloruro de guanidinio 2 M. El cloruro de guanidinio es un agente caotrópico con el que se han obtenido buenos resultados en la purificación de los reguladores transcripcionales FurA y FurB, en los que el uso de este compuesto en el tampón de purificación permitió purificarlos con un rendimiento alto y manteniendo su actividad de unión al DNA, y que se ha propuesto que podría permitir la purificación de otros reguladores transcripcionales diferentes a los de la familia FUR que presnten problemas de agregación (78). Además, a la vista de los patrones de elución (Figuras 20 y 21) se decidió purificar ambas proteínas empleando como ion metálico el cobalto; puesto que se aprecia como en la purificación con este ion metálico hay una menor presencia de contaminantes que cuando se empleó níquel.

Al igual que en la purificación con Tris 50 mM pH 8, en la purificación llevada a cabo con cloruro de guanidinio se sonicaron las células BL21 que sobreexpresaban las proteínas para liberar al exterior celular las proteínas de interés y se centrifugaron para obtener el extracto crudo, el cual se cargó en la columna de afinidad previamente cargada con cobalto. Posteriormente, se realizaron lavados para eliminar las proteínas unidas por interacciones inespecíficas y se eluyeron las proteínas de interés mediante un gradiente de imidazol (0-1 M). Al igual que en el caso anterior, se recogieron alícuotas en todos los pasos del proceso de purificación y estas alícuotas, junto a las fracciones obtenidas en la elución, se analizaron mediante electroforesis desnaturalizante en gel de poliacrilamida, previa precipitación con TCA. Los resultados de la electroforesis se muestran en la figura 22.



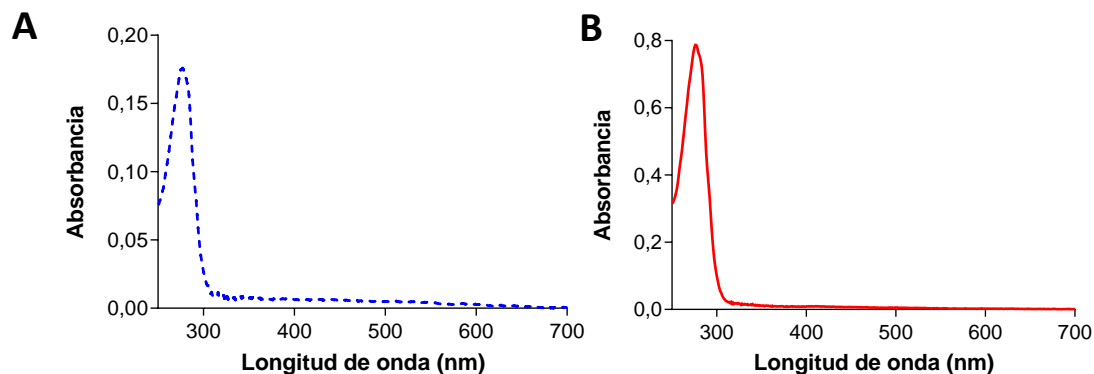
**Figura 22. Perfil de la elución de la columna de purificación de la proteína Alr1976 (A) y All0345 (B) empleando un tampón de purificación con cloruro de guanidinio 2 M y cobalto como ión metálico inmovilizado.** En ambos geles, el primer carril (MW) corresponde al marcador de pesos moleculares Low Molecular Weight Protein Marker II (NZY), el segundo carril (Son) corresponde a una alícuota recogida tras la sonicación, el tercero (Pel) al precipitado obtenido tras clarificar el sonicado, el cuarto (Ext) al extracto crudo, el quinto (FT) al volumen eluido al cargar el extracto crudo, el sexto (L1) al primer lavado, el séptimo (L3) al tercer lavado y los carriles restantes (F4-F48) a las fracciones obtenidas tras la elución de la columna con un gradiente de imidazol (0-1 M). El análisis se llevó a cabo mediante electroforesis desnaturalizante en gel de poliacrilamida al 17 %.

Como se muestra en la figura 22, tras la sonicación y centrifugación en tampón con cloruro de guanidinio, se consigue solubilizar la mayor parte de las proteínas, puesto que la banda que aparece en el carril del pellet (Pel) a una altura de 18,5 KDa en el caso de Alr1976 y a una altura de 40 KDa en el caso de All0345 es muy tenue, mientras que, en el carril del extracto (ext), esta banda es más gruesa. Por otro lado, se puede observar que, con los sucesivos lavados, se consigue eliminar de la columna las proteínas contaminantes, puesto que en el carril del tercer lavado (L3) apenas aparecen bandas de proteínas. En el caso de la proteína Alr1976 (figura 22A), se observa que esta comienza a eluir en la fracción 8 y se puede ver que en las primeras fracciones en las que hay proteína, especialmente en la 8 y en la 12, la concentración de esta es tan alta que incluso se deforman los carriles del gel. Se observa que el resto de la proteína eluye entre las fracciones 16 y 28 y que en las fracciones restantes (F32-F48), la concentración de Alr1976 es muy baja. En las fracciones 8, 12, 16 y 21 se observa que aparecen otras proteínas contaminantes de mayor peso molecular, aunque la proteína de interés se encuentra en mucha mayor proporción. En el caso de la proteína All0345 (figura 22B), se obtiene una banda a unos 40 KDa, correspondiente a la proteína All0345, en todas las fracciones analizadas. No obstante, es en las fracciones comprendidas entre la 8 y la 20 donde la concentración es mucho mayor, e incluso en la fracción 12 se puede ver que se deforman los carriles del gel.

Tras este análisis se seleccionaron las fracciones de interés y se llevó a cabo la diálisis de las proteínas. Para la proteína Alr1976 se combinaron las fracciones 6 a 21 y para All0345, las fracciones 16 a 21, que se introdujeron en membranas de diálisis y se sometieron 2 ciclos de diálisis a 4 °C en tampón acético-acetato a pH 4. Tras el primer ciclo de diálisis se observó que las proteínas precipitaban dentro de la membrana. En el segundo ciclo el regulador Alr1976 se resuspendió prácticamente por completo mientras que el regulador All0345, solo se resuspendió parcialmente, por lo que se centrifugó para eliminar la proteína precipitada.



Finalmente, se analizaron las proteínas mediante espectroscopía UV-Visible (Figura 23) y se cuantificaron haciendo uso de la absorbancia a 280 nm, del coeficiente de extinción teórico y de la ley de Beer-Lambert. A diferencia del protocolo anterior, la purificación con cloruro de guanidinio permitió obtener las proteínas a una alta concentración. En el caso de Alr1976, se obtuvo a 390  $\mu\text{M}$  y para All0345, aunque parte precipitó en la diálisis, se obtuvo a 80  $\mu\text{M}$ .

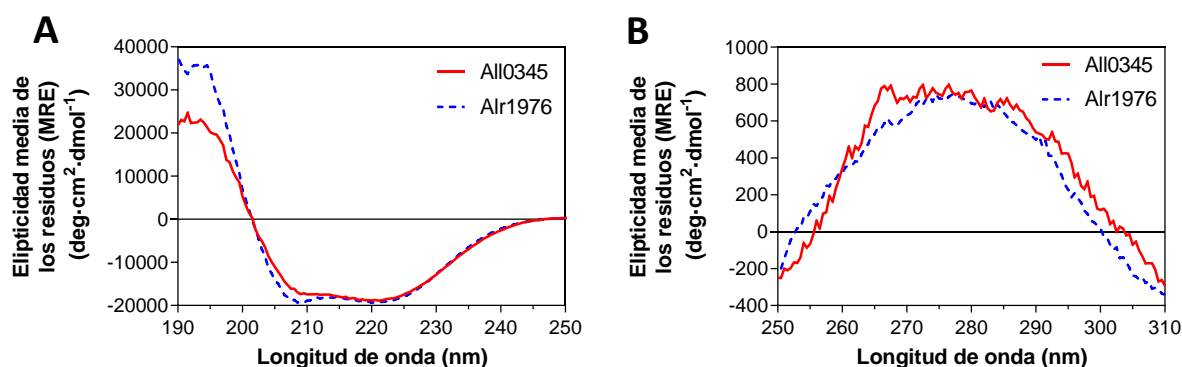


**Figura 23.** Espectros UV-Visible entre 250 y 700 nm de soluciones 20  $\mu\text{M}$  de los reguladores Alr1976 (A) y All0345 (B) tras la purificación con cloruro de guanidinio y diálisis en acético-acetato 10 mM pH 4

Como se observa en la Figura 23, los espectros UV-Visible de los reguladores Alr1976 y All0345 no presentan picos en la región del visible, lo que indica la ausencia de grupos prostéticos con absorbancia en esta región, y aparece un pico a 280 nm, resultado de la absorción de los aminoácidos aromáticos de las proteínas.

### 5.6. Análisis de la estructura secundaria y terciaria mediante dicroísmo circular

Tras haber purificado las proteínas de interés se llevó a cabo un análisis de las mismas mediante dicroísmo circular en el UV-lejano, que proporciona información sobre la estructura secundaria, y en el UV-cercano, que proporciona información sobre la estructura terciaria. Los resultados se muestran en la figura 24.



**Figura 24.** Espectros de dicroísmo circular UV-lejano (A) y UV-cercano (B) de soluciones 10  $\mu\text{M}$  de los reguladores Alr1976 y All0345.

En el UV-lejano, la absorción es dependiente de los ángulos  $\phi$  y  $\psi$  de los enlaces peptídicos, los cuales son diferentes dependiendo del tipo de estructura secundaria. Esto ocasiona que cada motivo de estructura secundaria tenga un espectro característico: las hélices  $\alpha$  tienen dos mínimos en torno a 222 y 208 nm y un máximo en torno a 190 nm mientras que las hojas  $\beta$  presentan un mínimo a 215-218 nm y un máximo a 196-198 nm.

Como se aprecia en la figura 24 A, ambas proteínas son ricas en hélices  $\alpha$ , sobre todo Alr1976. La deconvolución de los espectros permitió estimar los porcentajes de cada tipo de estructura para las dos proteínas. Se estima que Alr1976 está constituida por un 56,5 % de hélice  $\alpha$  y un 0,6 % de lámina  $\beta$ , lo que se asemeja con la predicción de estructura secundaria (Figura 6), en la que se predice un 66 % de hélice  $\alpha$ . Para All0345, se estima que se compone de un 40% de hélice  $\alpha$  y un 17 % de lámina  $\beta$ , que también es semejante a la predicción de estructura secundaria, que presentaba un 41 % de hélice  $\alpha$  y 22 % de lámina  $\beta$  (Figura 6).

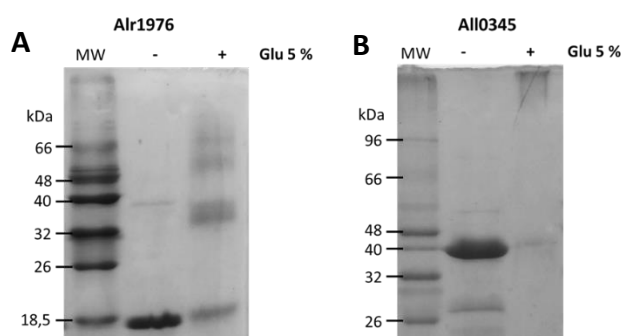
En el UV-cercano absorben los aminoácidos aromáticos, en concreto el triptófano, pero únicamente cuando se encuentran en el interior de una estructura, de modo que, en las proteínas desplegadas, el triptófano no presenta absorción. Por tanto, esta técnica permite estudiar si la proteína se encuentra plegada o desplegada. Como se aprecia en la figura 24B, ambas proteínas dan un espectro UV-cercano, lo que indica que se han obtenido plegadas en la purificación.

### 5.7. Determinación de la presencia de metal por ICP-OES

Se determinó mediante ICP-OES la presencia de metales característicos de las metaloproteínas (Cd, Co, Cu, Fe, Mg, Mn, Ni, Pb, Zn y Mo) en las proteínas Alr1976 y All0345. En todos los casos, los metales estaban por debajo del límite de detección, lo que indica que las proteínas Alr1976 y All0345 no se han purificado conteniendo metales.

### 5.8. Análisis de oligomerización por entrecruzamiento

Dado que los estudios bioinformáticos sugerían que las proteínas Alr1976 y All0345 podían dimerizar, se analizó su oligomerización mediante ensayos de entrecruzamiento con glutaraldehído al 5 %, un entrecruzador bifuncional. Los resultados se muestran en la figura 25.



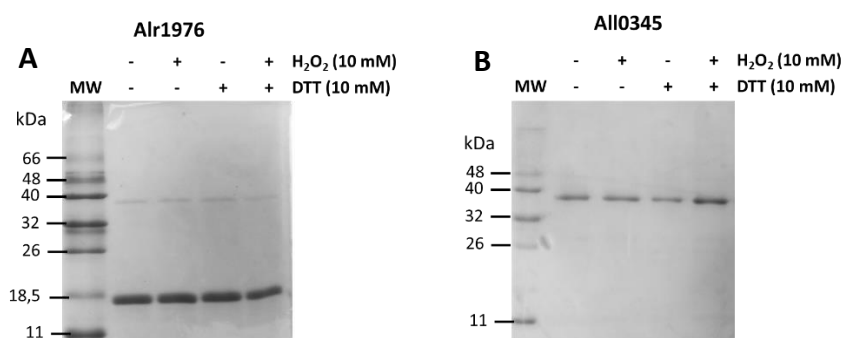
**Figura 25. Ensayos de oligomerización por entrecruzamiento de las proteínas Alr1976 (A) y All0345 (B) con glutaraldehído al 5 %.** En ambos geles, el primer carril (MW) corresponde al marcador de pesos moleculares Low Molecular Weight Protein Marker II (NZY), el segundo carril (Son) corresponde a la proteína (10  $\mu$ M) sin agente de entrecruzamiento y el tercer carril corresponde a la proteína (10  $\mu$ M) tratada con glutaraldehído 5 % (Glu 5%). El análisis se llevó a cabo mediante electroforesis desnaturalizante en gel de poliacrilamida al 15 %.

En la figura 25A, en el segundo carril se aprecia una banda en torno a 18,5 KDa, que corresponde a Alr1976, y en el tercer carril, se aprecia una disminución de la intensidad de dicha banda, mientras que aparecen una banda entre 32 y 40 kDa, que podría corresponder a un dímero de Alr1976, que tendría un peso teórico de 26,8 KDa, y unas bandas por encima de 48 KDa, que podrían corresponder a oligómeros de este regulador. Por otro lado, para All0345, en la figura 25B, se aprecia como en el carril tratado con glutaraldehído desaparece la banda de en torno a 40 KDa, y aparece una banda con un peso molecular superior a 96 KDa, que podría corresponder a agregados de la proteína u oligómeros de alto peso molecular.



### 5.9. Análisis de oligomerización mediada por puentes disulfuro

Dado que en ambos casos se observó la formación de oligómeros de un tamaño molecular superior y ambas proteínas contenían cisteínas se decidió estudiar si esta oligomerización podía estar mediada por puentes disulfuro. Para ello, se llevó a cabo un estudio de oligomerización tratando las proteínas con  $H_2O_2$ , un agente oxidante que produce la formación de puentes disulfuro y un agente reductor (DTT) que reduce los puentes disulfuro. El resultado de este análisis se muestra en la figura 26.

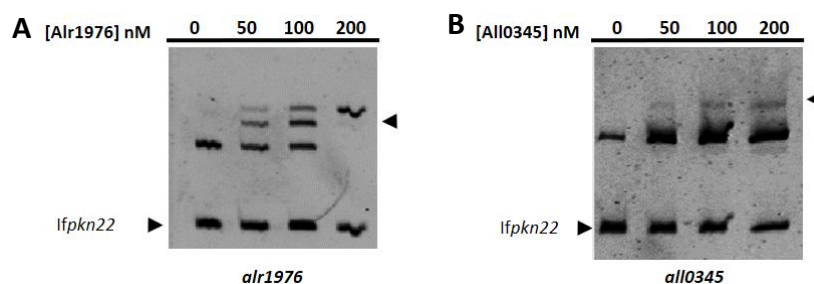


**Figura 26. Ensayos de oligomerización mediados por puentes disulfuro de las proteínas Alr1976 (A) y All0345 (B).** En ambos geles, el primer carril (MW) corresponde al marcador de pesos moleculares Low Molecular Weight Protein Marker II (NZY), el segundo carril corresponde a la proteína sin tratar, el tercer carril a la proteína tratada durante 10 min con  $H_2O_2$  10 mM, el cuarto carril a la proteína tratada durante 10 min con DTT 10 mM y el quinto a la proteína tratada durante 10 min con  $H_2O_2$  10 mM y después durante 10 min con DTT 10 mM. El análisis se llevó a cabo mediante electroforesis desnaturalizante en gel de poliacrilamida al 15 %.

En la figura 26 A, se puede apreciar como en los carriles peróxido de hidrógeno y de DTT, aparecen las mismas bandas que en el carril control (carril 1): una a 18,5 KDa, correspondiente a Alr1976 y otra en torno a los 40 KDa, correspondiente a una contaminación. De la misma manera, en el caso de All0345, se observa en la figura 26 B, que la adición de DTT y peróxido de hidrógeno aparentemente no afectan a la proteína, pues en todos los carriles se observa únicamente una banda a 40 KDa correspondiente a la proteína All0345.

### 5.10. Determinación de las condiciones de unión a DNA mediante EMSA

Para determinar si los reguladores Alr1976 y All0345 presentaban actividad de unión al DNA se llevaron a cabo ensayos de retardo en gel utilizando como diana su propia región promotora, pues la mayoría de reguladores transcripcionales están autorregulados (Figura 27).

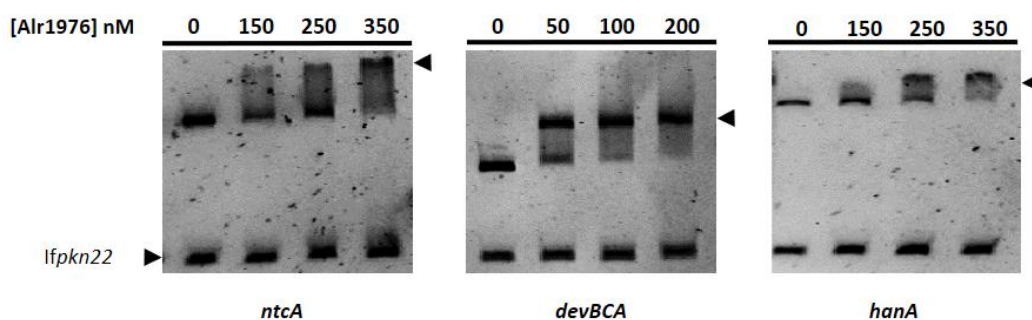


**Figura 27. Ensayos de retardo en gel de las proteínas Alr1976 y All0345 frente a sus respectivos promotores. A.** EMSA de Alr1976. El primer carril corresponde al promotor libre en ausencia de proteína y en los siguientes tres carriles, dicho promotor con concentraciones crecientes de Alr1976: 50, 100, 200 y 400 nM. **B.** EMSA de All0345. El carril 1, corresponde al promotor libre en ausencia de proteína y los siguientes a dicho promotor incubado con concentraciones crecientes de All0345: 0, 25, 50, 100, 150, 200 y 250 nM. En ambos casos, los ensayos se realizaron en presencia de DNA inespecífico (*ifpkn22*) y los promotores libres o incubados con proteína se corrieron en geles de poliacrilamida del 6 %.

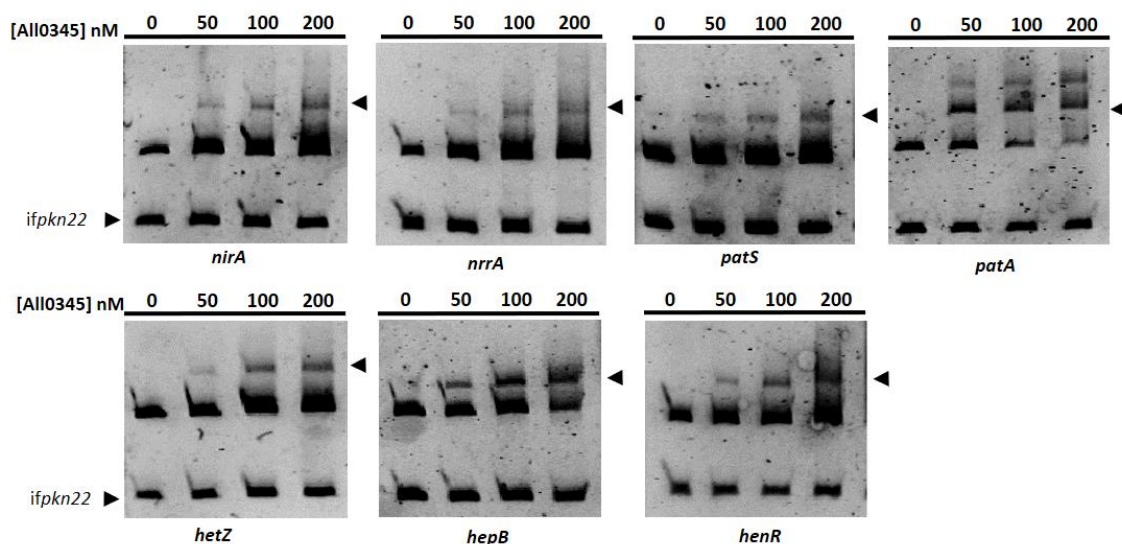
Como se aprecia en la figura 27, tanto Alr1976 como All0345 tienen actividad de unión a DNA y son capaces de unirse a su propio promotor de forma específica, pues en ambos casos se forman bandas correspondientes a complejos proteína DNA con la región promotora y no se produce retardo del DNA inespecífico. Además, ambos reguladores son capaces de llevar a cabo dicha unión sin necesidad de añadir metales ni agentes reductores (DTT).

### 5.11. Estudio de la unión a regiones promotoras de genes implicados en el metabolismo del nitrógeno y la diferenciación de heterocistos.

Debido a los indicios de la posible implicación de las proteínas Alr1976 y All0345 en el metabolismo del nitrógeno y formación de heterocistos, pues la expresión de *alr1976* disminuye 2,5 veces tras 12 horas en deficiencia de nitrógeno (67) y *all0345* muestra expresión diferencial en el transcriptoma de una estirpe de sobreexpresión de FurC en *Anabaena* únicamente en condiciones de deficiencia de nitrógeno (64), se analizó mediante ensayos de retardo en gel si estas proteínas se unían a los promotores de los principales genes implicados en estos procesos. Se estudiaron genes implicados en la asimilación de nitrógeno como *nirA*, *nifH* y *glnA*, genes de las primeras etapas de la diferenciación del heterocisto como *hetR*, *ntcA*, *nrrA* y *hanA*; genes implicados en la formación del patrón como *asr1734*, *patA*, *patS*, *hetN*, *hetC* y *hetZ*; y genes implicados en la formación de las envueltas del heterocisto como *henR*, *hepA*, *hepB*, *hepC*, *hepK*, *devH*, *devBCA* y *hglE*. Para obtener las correspondientes regiones promotoras, se amplificaron por PCR fragmentos de DNA de entre 300 y 400 pb localizados upstream del sitio de inicio de la transcripción de estos genes. Los ensayos de retardo en gel en los que se obtuvo que los reguladores Alr1976 y All0345 se unían a la región promotora se muestran en las figuras 28 y 29.



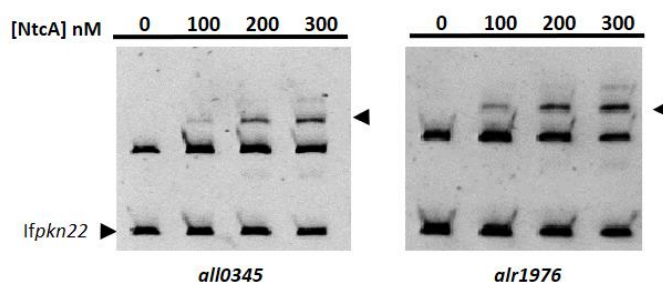
**Figura 28. Ensayos de retardo en gel de Alr1976 con las regiones promotoras de los genes *ntcA*, *devBCA* y *hanA*.** El carril 1 corresponde al promotor libre (en ausencia de proteína) y los carriles siguientes corresponden al promotor incubado con concentraciones crecientes de proteína: 150, 250 y 350 nM en el caso de *ntcA* y *hanA*; y 50, 100 y 200 nM en el caso de *devBCA*. Todos los ensayos se llevaron a cabo en presencia de DNA inespecífico (*lfpkn22*) y el promotor libre o incubado con proteína se analizó en geles de poliacrilamida al 6 %.



**Figura 29.** Ensayos de retardo en gel de All0345 con las regiones promotoras de genes implicados en el metabolismo del nitrógeno y formación de heterocistos. El carril 1 corresponde al promotor libre (en ausencia de proteína) y los carriles siguientes corresponden al promotor incubado con concentraciones crecientes de proteína: 50, 100 y 200 nM. Todos los ensayos se llevaron a cabo en presencia de DNA inespecífico (*ifpkn22*) y el promotor libre o incubado con proteína se analizó en geles de poliacrilamida al 6 %.

La figura 28 muestra que Alr1976 es capaz de unirse a las regiones promotoras de *ntcA*, *devBCA* y *hanA*, puesto que se observa un retardo de la banda del promotor al aumentar las concentraciones de Alr1976. Por otro lado, en la figura 29, se aprecia como All0345 se une a los promotores de *nirA*, *hetZ*, *patS*, *henR*, *patA*, *hepB* y *nrrA*. Se observa en todos los casos, una banda superior al promotor de interés que constituye el complejo formado por la proteína y el DNA, y que al tener un mayor peso molecular tiene una migración retardada.

Finalmente, dado que NtcA es el principal regulador del metabolismo del nitrógeno, que Alr1976 y All0345 se unen a promotores de genes que participan en esta vía y además NtcA regula muchos de los genes modulados por las proteínas FUR se decidió estudiar si NtcA se unía a los promotores de los genes *alr1976* y *all0345*. Los resultados de los ensayos de retardo en gel se muestran en la figura 30.



**Figura 30.** Ensayos de retardo en gel de NtcA frente a las regiones promotoras de *all0345* y *alr1976*. El carril 1 corresponde al promotor libre (en ausencia de proteína) y los carriles siguientes corresponden al promotor incubado con concentraciones crecientes de proteína: 100, 200 y 300 nM. Todos los ensayos se llevaron a cabo en condiciones reductoras (con DTT) y en presencia de DNA inespecífico (*ifpkn22*) y el promotor libre o incubado con proteína se analizó en geles de poliacrilamida al 6 %.

Como se observa en la figura 30, NtcA se une a las regiones promotoras de *alr1976* y *all0345*, ya que se aprecia una banda de un peso molecular mayor al promotor libre, correspondiente al complejo NtcA-promotor.

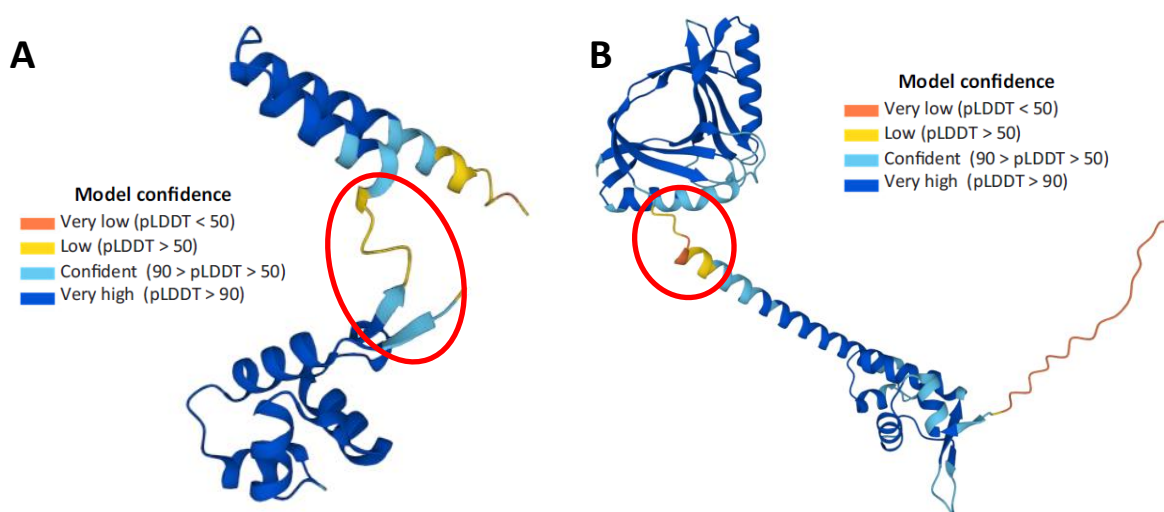
## 6. Discusión

### 6.1. Análisis bioinformático de los reguladores Alr1976 y All0345

En este trabajo se han purificado y estudiado dos reguladores transcripcionales de la cianobacteria *Anabaena* que forman parte de la red de regulación de FurC, Alr0345 y Alr1976. El análisis bioinformático ha aportado información relevante tanto para diseñar las condiciones de su purificación como sus potenciales funciones.

La caracterización bioinformática de Alr1976 mostró que este regulador pertenece a la familia de reguladores transcripcionales XRE (figura 5A) y que además comparte los motivos de estructura secundaria característicos de esta familia (figura 6A). Al construir los modelos de estructura terciaria se observó que las predicciones de Swiss-Model era diferente a la de AlphaFold. A diferencia de los primeros, este último presenta una pequeña lámina  $\beta$ , tal y como se aprecia en la figura 8A. La predicción de estructura secundaria (figura 6A), así como la información bibliográfica sobre las proteínas de la familia XRE (68, 69, 73) sugiere que Alr1976 contiene únicamente hélices  $\alpha$ , por lo que es altamente probable que los modelos por homología de Swiss-Model sean más cercanos a la estructura real que el de AlphaFold. Además, en la figura 31A se observa que el nivel de confianza de los residuos de la proteína que formarían la lámina  $\beta$  predicha por AlphaFold están junto a una zona desordenada y el grado de confianza, a diferencia de la mayor parte de la estructura, es menor.

El análisis filogenético (figura 14) mostró que, de los dos reguladores utilizados como molde para construir la estructura terciaria (BswR de *P. aeruginosa* y EspR de *M. tuberculosis*), el regulador BswR es el más cercano filogenéticamente, de modo que posiblemente la estructura terciaria de Alr1976 sea más semejante a la de este regulador que a la de EspR. Al modelar la interacción con el DNA se observó que mientras que en el modelo construido en base al regulador EspR se predecía que la unidad funcional que interaccionaba con el DNA era un dímero de dímeros (figura 10B) (73), en la predicción en base al regulador BswR era un dímero (figura 10A) cuya formación estaba mediada por interacciones hidrofóbicas entre residuos hidrofóbicos de la hélice  $\alpha$  del extremo C-terminal (figura S-1 A). Dado que Alr1976 es filogenéticamente más cercano a este regulador, es posible que su unión al DNA sea también similar a la de este regulador. No obstante, el porcentaje de identidad de estas proteínas con Alr1976 no es demasiado elevado, por lo que estos resultados únicamente son orientativos.



**Figura 31.** Nivel de confianza en la predicción de estructura terciaria de Alr1976 (A) y All0345 (B) por AlphaFold. La confianza de cada región se representa con una escala de color comprendida entre 0 y 100.

Con respecto a All0345, los estudios bioinformáticos muestran que pertenece a la familia de reguladores transcripcionales MerR (figura 5B). Al igual que los reguladores de esta familia presenta un dominio de unión a DNA en el extremo amino-terminal, una larga hélice  $\alpha$  central implicada en la dimerización y un dominio de unión a efector en el extremo C-terminal (figura 7B) (71). En este caso las predicciones de estructura terciaria de Swiss-Model y AlphaFold fueron similares, salvo para el bucle que conecta la hélice  $\alpha$  implicada en la dimerización con el dominio efector del extremo C-terminal, lo que hacía que la orientación de este dominio variase entre los modelos (figura 8B). En la figura 31B se observa que la confianza en la parte final de la hélice  $\alpha$  implicada en la dimerización, así como el bucle que la conecta con el dominio de unión a efector es baja, lo que podría explicar que la predicción de la orientación del dominio de unión a efector difiera entre los modelos. En el caso de los modelos en complejo con DNA (figuras 10C y 10D), se observa que se produce una deformación del DNA, la cual se ha descrito previamente para los reguladores MerR, que activan la transcripción deformando el DNA de forma que los elementos -10 y -35 quedan en una conformación óptima para interactuar con la RNA polimerasa (76).

Dentro de la familia MerR hay una gran diversidad en el dominio de unión a efector del extremo C-terminal, pues los miembros de esta familia son capaces de detectar una gran variedad de compuestos (72). Algunos de ellos son capaces de unirse a cationes metálicos como MerR de *S. aureus* o CueR de *E. coli*, otros pueden detectar estrés oxidativo como SoxR de *E. coli* y hay otros que son capaces de unir compuestos químicos orgánicos, incluyendo diversos fármacos, como BmrR y BltR de *B. subtilis* y EcmrR de *E. coli* (72). All0345 presenta una mayor cercanía filogenética con estos últimos (figura 15), lo que sumado a la presencia de dos bolsillos hidrofóbicos que podrían unir compuestos químicos (figura 11), sugiere que All0345 pertenece a la subfamilia de unión a compuestos orgánicos. Por tanto, futuros estudios irán encaminados al estudio de la interacción de All0345 con diferentes metabolitos que pudieran actuar como señalizadores en la cianobacteria y posiblemente modular la actividad de unión al DNA de All0345.

Como se observa en las figuras 12 y 13, los homólogos de las proteínas Alr1976 y All0345 además de estar presentes en cianobacterias fijadoras de nitrógeno como *Tolypothrix* sp. y *Trichormus* sp., también lo están en estirpes no fijadoras, como *Chamaesiphon* sp. y *Gloeocapsa* sp. Por ello, aunque en este trabajo se ha centrado en el estudio de su papel en el control de genes implicados en metabolismo de nitrógeno y formación de heterocistos, es posible que estos reguladores participen también en otros procesos celulares.

## 6.2. Purificación y caracterización de los reguladores Alr1976 y All0345

La sobreexpresión de las proteínas Alr1976 y All0345 en *E. coli* BL21 se produjo en todas las condiciones estudiadas de forma satisfactoria. No obstante, las proteínas mostraron una baja solubilidad en las condiciones de diálisis a pH 8 y precipitaron, obteniendo concentraciones insuficientes de proteína para llevar a cabo estudios posteriores. Por ello, se llevó a cabo una purificación empleando cloruro de guanidinio, que previamente se había utilizado con buenos resultados para la purificación de FurA (78) y que en este caso permitió recuperar altas concentraciones de las proteínas Alr1976 y All0345 tras la diálisis. Además, ambas proteínas se obtuvieron plegadas, puesto que el dicroísmo circular en la región del UV cercano mostró absorción (figura 24B) y activas, pues eran capaces de unirse a sus propios promotores (figura 27). Esto corrobora que este protocolo de purificación es eficaz para otros reguladores transcripcionales diferentes de las proteínas Fur, como se había propuesto previamente (78).



Respecto a la caracterización bioquímica, los resultados de la determinación de metales indican que las proteínas Alr1976 y All0345 no poseen metales estructurales, lo que concuerda con las estructuras resueltas de los reguladores de la familia XRE (69, 73) y de los reguladores tipo MerR (72, 76) empleadas como modelos en este trabajo, las cuales no contienen metales estructurales.

Los ensayos de entrecruzamiento muestran que Alr1976 forma tanto dímeros como oligómeros de mayor tamaño (figura 25A), cuya formación no está mediada por puentes disulfuro (figura 26A). Esto concuerda con lo que sucede en los reguladores de la familia XRE, en los que generalmente, la dimerización tiene lugar debido a interacciones entre residuos hidrofóbicos presentes en la hélice de dimerización (68). Por el contrario, en el caso de All0345 no se observó la formación de dímeros en los ensayos de entrecruzamiento (figura 25B), lo que podría deberse a que el glutaraldehído ocasione la agregación inespecífica de la proteína y su precipitación, dando lugar a formas de alto peso molecular. Para solventar este problema en el futuro se evaluará la potencial formación de dímeros u otros oligómeros mediante otras técnicas como MALLS (Multi-Angle Laser Light Scattering). No obstante, para comprobar si la dimerización de los reguladores ocurre *in vivo* serían necesarios ensayos de doble híbrido, en los que es posible el estudio de interacciones proteína-proteína en condiciones fisiológicas.

Tanto Alr1976 como All0345 se unen a su propio promotor (figura 27), indicando que son capaces de autorregularse como sucede habitualmente en los reguladores de las familias XRE y MerR, respectivamente (68, 70). Además, en base a los resultados de este trabajo, no requieren metales ni condiciones reductoras para unirse al DNA (figura 27). Alr1976, perteneciente a la familia XRE, contiene únicamente el dominio de unión a DNA y una hélice de dimerización, proceso llevado a cabo por interacciones hidrofóbicas, que tampoco está modulado por la presencia de metales ni por el estado redox (68). Por lo tanto, serán necesarios estudios en el futuro para identificar las señales que modulan su interacción con el DNA.

Los reguladores transcripcionales de la familia MerR, familia a la que pertenece All0345, presentan un dominio C-terminal muy divergente (70). En el caso de All0345, además de la similitud con reguladores de la familia MerR capaces de unir compuestos químicos (figura 15) y de la existencia de dos bolsillos hidrofóbicos (figura 11), otra evidencia a favor de que All0345 pertenece a la subfamilia de los reguladores MerR cuya unión al DNA está modulada por la interacción con ligandos es su capacidad de unión al DNA en ausencia de metales y en ausencia de agente reductor, pues si perteneciera a las subfamilias de unión a metales o de respuesta oxidativa cabría esperar que su unión al DNA se viera afectada por estos factores. Por ello, en trabajos posteriores sería de interés analizar la unión de All0345 al DNA en presencia de diferentes compuestos químicos, que pudieran actuar como efectores.

### 6.3. Participación de Alr1976 y All0345 en la regulación del metabolismo del nitrógeno y la diferenciación de heterocistos en *Anabaena*

Dada la importancia del metabolismo del nitrógeno en cianobacterias y la posible implicación de los reguladores Alr1976 y All0345 en este proceso, ya se observan cambios en la expresión de los genes *alr1976* y *all0345* en diferentes transcriptomas en deficiencia de nitrógeno, se estudió su posible implicación en la regulación de genes implicados en este proceso.

Como se observa en la figura 28, el regulador Alr1976 se une a los promotores de los genes *hanA*, *ntcA* y *devBCA*. HanA es esencial para la diferenciación de heterocistos en *Anabaena* (39), actuando principalmente en las primeras etapas de diferenciación y favoreciendo la expresión de *hetR* (79) mientras que *devBCA* es un operón formado por los genes *devB*, *devC* y *devA*, los cuales codifican para los componentes de un transportador tipo ABC (47). Concretamente, DevBCA funciona como un exportador que transloca glicolípidos específicos de membrana y es esencial para la formación de la capa glicolípida de los heterocistos (80) y su expresión es dependiente de HetR, aunque NtcA también juega un papel esencial como activador del operón *devBCA* (81). Estudios transcripcionales mostraron que la expresión de *devBCA* y de *hanA* aumenta a las 12 horas tras la deficiencia de nitrógeno, mientras que en este momento la expresión de *alr1976* disminuye (67). Por tanto, teniendo en cuenta que los reguladores de la familia XRE, a la que pertenece Alr1976 suelen funcionar como represores (68), es posible que Alr1976 se una a los promotores del operón *devBCA* y *hanA* reprimiendo su expresión y que al disminuir la expresión de *alr1976* tras 12 horas en deficiencia de nitrógeno se desrepriman los genes *devBCA* y *hanA* y aumente su expresión.

Por otro lado, *alr1976* está regulado por NtcA (figura 30) y además regula a *ntcA*, puesto que se une a su promotor (figura 28), generando así un ciclo. Dado que la mayoría de los reguladores de la familia XRE actúan como represores (68), Alr1976 podría actuar controlando la expresión de *ntcA* a lo largo del proceso de diferenciación de heterocistos, reprimiendo su expresión en el momento adecuado para mantener unos niveles adecuados de NtcA que permitan una correcta diferenciación a heterocistos.

Respecto a All0345, como se aprecia en la figura 29 fue capaz de unirse a los promotores del operón *nir-nrt-nar*, implicado en la asimilación de nitrato, así como de los genes *nrrA*, que actúa en las primeras etapas de diferenciación del heterocisto; *hetZ*, *patS* y *patA*, implicados en la formación del patrón; y *henR* y *hepB*, que participan en la síntesis de las envueltas del heterocisto.

El operón *nir* está formado por la nitrito reductasa (*nirA*), el transportador de tipo ABC de nitrito/nitrato (*nrtABCD*) y la nitrato reductasa (*narB*) (24) y su expresión está controlada por el regulador NtcA (82), así como por FurC (64). Además, All0345 está regulado por FurC y en este trabajo se ha descrito que NtcA se une a su promotor (figura 30). Por ello se puede proponer que All0345 colabore en la activación de este operón de forma conjunta con FurC y NtcA.

NrrA es un regulador de respuesta que facilita la diferenciación de heterocistos y permite que NtcA active la expresión de *hetR* de manera indirecta (83). All0345 se une al promotor de *nrrA* y NtcA se une además al promotor de *all0345* (figura 30), de forma que podría constituir una vía adicional para regular la expresión de *hetR* durante la diferenciación de heterocistos.

All0345 se une además a tres promotores de genes implicados en la formación del patrón de heterocistos: *patA*, *patS* y *hetZ*. PatS es un regulador negativo de HetR, el cual interacciona con este impidiendo su unión al DNA y por tanto la diferenciación de heterocistos (84), mientras que *patA* tiene una influencia en la diferenciación de heterocistos, promoviendo la expresión de *hetR*

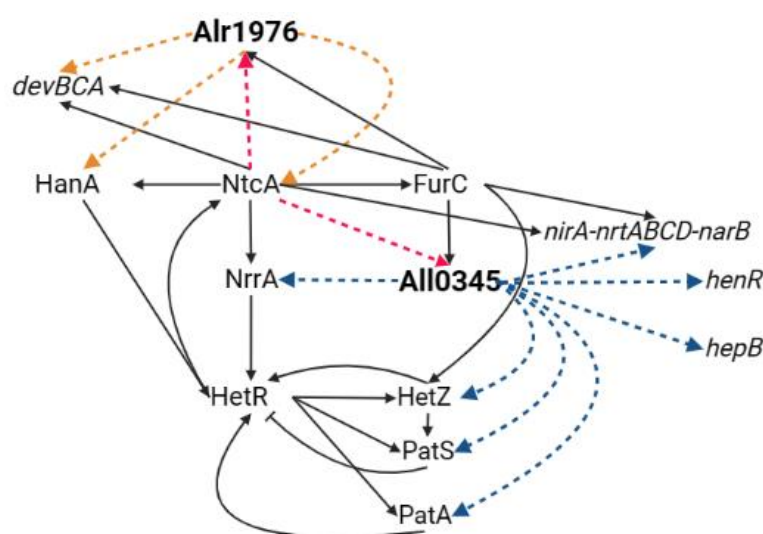


(85). HetZ, al igual que PatA favorece la expresión de *hetR*, pero además favorece también la de *patS* aumentando de esta forma el ciclo que genera el patrón. Además *patA*, *patS* y *hetZ* son activados por HetR (45). All0345 se une a los promotores de estos tres genes, por lo que All0345 podría jugar un papel en la formación del patrón del heterocisto. Por otro lado, se ha descrito que FurC regula de manera indirecta a *patA*, pues la sobreexpresión de FurC, altera su expresión; pero FurC no es capaz de unirse a su promotor (64). Dado que *all0345* está regulado por FurC y All0345 se une al promotor de *patA*, es posible que la regulación de la expresión de *patA* por parte de FurC transcurra de forma indirecta a través del regulador All0345.

Finalmente, All0345 también participa en la regulación de la síntesis de las envueltas, tanto en la formación de la capa Hep, pues se une al promotor de *hepB*, una glicosiltransferasa que participa en la formación de esta capa y cuya ausencia impide su síntesis (86), como en la formación de la capa Hgl, ya que se une al promotor de *henR*, el cual es un regulador de respuesta implicado no solo en la regulación de la síntesis de la capa Hep, sino también en de la capa Hgl (87).

De esta forma, All0345 está implicado en casi todas las fases de diferenciación de heterocistos, ya que regula a *nrrA*, de la fase de iniciación, a *patS* y *patA* de la formación del patrón, a *hetZ*, que se encuentra entre la fase de formación del patrón y la de compromiso y por último en la fase de morfogénesis, regulando a *hepB* y *henR*. Todos estos resultados se integran en la figura 32, donde se muestra un esquema de la red de regulación de genes del metabolismo del nitrógeno que incluye a los reguladores Alr1976 y All0345.

En base a los resultados de este trabajo, futuros estudios irán encaminados a la caracterización de mutantes de delección de estos reguladores para conocer si son activadores o represores de los genes a cuyos promotores se unen, además de verificar si estos genes son dianas directas de los reguladores Alr1976 y All0345. Por otro lado, también sería necesaria la realización de estudios *in vivo*, incluyendo una caracterización detallada del fenotipo de las estirpes mutantes, tanto en condiciones normales como en deficiencia de nitrógeno, para conocer en detalle las funciones fisiológicas de estos reguladores.



**Figura 32. Esquema de la red de regulación en la que participan los reguladores Alr1976 y All0345.** En líneas discontinuas se representan las relaciones descubiertas en este trabajo; en amarillo las de Alr1976, en azul las de All0345 y en rosa las de NtcA. Las líneas negras continuas indican relaciones ya conocidas, las acabadas en punta indican activación, mientras que las acabadas en barra indican inhibición.

## 7. Conclusiones

Tras la realización de este Trabajo de Fin de Máster se obtienen las siguientes conclusiones:

- Los estudios bioinformáticos permiten concluir que Alr1976 y All0345 son reguladores transcripcionales de la familia XRE y MerR respectivamente, y que All0345 posee un dominio de unión a efector potencialmente implicado en la unión de ligandos.
- El protocolo de purificación basado en el uso de una cromatografía de afinidad a iones metálicos inmovilizados (IMAC) con cobalto como ion metálico y de un tampón de purificación con cloruro de guanidinio 2 M permite obtener las proteínas Alr1976 y All0345 en altas concentraciones, plegadas y con actividad de unión a DNA.
- Las proteínas Alr1976 y All0345 son capaces de unirse *in vitro* al DNA de forma específica y sin necesidad de condiciones reductoras o presencia de metales.
- Alr1976 está implicada en la modulación del metabolismo del nitrógeno, puesto que se une al promotor de *ntcA*, y en el control de genes implicados en la diferenciación del heterocisto, que participan tanto en las etapas tempranas de la diferenciación como en la síntesis de la envuelta.
- All0345 participa en la regulación de la asimilación de nitrato, así como en las etapas de iniciación, formación del patrón y síntesis de las envueltas durante el proceso de diferenciación de heterocistos.

## Conclusions

After finishing this Final Master Project, the following conclusions have been obtained:

- Bioinformatic studies revealed that Alr1976 and All0345 are transcriptional regulators from the XRE and MerR family, respectively, and that All0345 contains an effector binding domain possibly involved in ligand binding.
- The purification protocol based on the use of an immobilized metal ion affinity chromatography (IMAC) with cobalt as metallic ion and a purification buffer with guanidine hydrochloride 2 M yields Alr1976 and All0345 proteins in high concentrations, folded and with DNA binding activity.
- Alr1976 and All0345 proteins are able to bind to DNA *in vitro* in a specific manner, and without the requirement of reducing conditions or the presence of metals.
- Alr1976 is involved in the modulation of nitrogen metabolism, since it binds to the *ntcA* promoter and in the control of genes involved in heterocyst differentiation both in the early stages of differentiation and in envelope biosynthesis.
- All0345 is involved in the modulation of nitrate assimilation, as well as in the stages of initiation, pattern formation and envelope synthesis during heterocyst differentiation.

## 8. Bibliografía

1. Steuer R, Knoop H, Machné R. Modelling cyanobacteria: from metabolism to integrative models of phototrophic growth. *J Exp Bot.* 2012;63(6):2259-74.
2. Mehdizadeh Allaf M, Peerhossaini H. Cyanobacteria: Model Microorganisms and Beyond. *Microorganisms.* 2022;10(4).
3. Rachedi R, Foglino M, Latifi A. Stress Signaling in Cyanobacteria: A Mechanistic Overview. *Life.* 2020;10(12).
4. Ho MY, Soulier NT, Canniffe DP, Shen G, Bryant DA. Light regulation of pigment and photosystem biosynthesis in cyanobacteria. *Curr Opin Plant Biol.* 2017;37:24-33.
5. Stadnichuk IN, Kusnetsov VV. Phycobilisomes and Phycobiliproteins in the Pigment Apparatus of Oxygenic Photosynthetic: From Cyanobacteria to Tertiary Endosymbiosis. *Int J Mol Sci.* 2023;24(3).
6. Vidal L, Ballot A, Azevedo S, Padisak J, Welker M. Introduction to cyanobacteria. 2021. p. 163-211.
7. Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Microbiology.* 1979;111(1):1-61.
8. Sánchez-Baracaldo P, Cardona T. On the origin of oxygenic photosynthesis and Cyanobacteria. *New Phytol.* 2020;225(4):1440-6.
9. Knoll A. Cyanobacteria and Earth History. En: Herrero A, Flores E, editores. *The cyanobacteria: Molecular Biology, Genomics, and Evolution.* 2008.
10. Demoulin CF, Lara YJ, Cornet L, François C, Baurain D, Wilmotte A, *et al.* Cyanobacteria evolution: Insight from the fossil record. *Free Radical Bio Med.* 2019;140:206-23.
11. Sánchez-Baracaldo P, Bianchini G, Wilson JD, Knoll AH. Cyanobacteria and biogeochemical cycles through Earth history. *Trends Microbiol.* 2022;30(2):143-57.
12. Huisman J, Codd GA, Paerl HW, Ibelings BW, Verspagen JMH, Visser PM. Cyanobacterial blooms. *Nat Rev Microbiol.* 2018;16(8):471-83.
13. Herrero A, Stavans J, Flores E. The multicellular nature of filamentous heterocyst-forming cyanobacteria. *FEMS Microbiol Rev.* 2016;40(6):831-54.
14. Meeks John C, Elhai J. Regulation of Cellular Differentiation in Filamentous Cyanobacteria in Free-Living and Plant-Associated Symbiotic Growth States. *Microbiol Mol Biol R.* 2002;66(1):94-121.
15. Stal LJ. Nitrogen Fixation in Cyanobacteria. En: Sons JW, editor. *Encyclopedia of Life Sciences.* 2015. p. 1-9.
16. Kumar K, Mella-Herrera RA, Golden JW. Cyanobacterial Heterocysts. *CSH Perspect Biol.* 2010;2(4).
17. Flores E, Herrero A. Compartmentalized function through cell differentiation in filamentous cyanobacteria. *Nat Rev Microbiol.* 2010;8(1):39-50.
18. Vijayakumar S, Menakha M. Pharmaceutical applications of cyanobacteria—A review. *J Acute Med.* 2015;5(1):15-23.
19. Singh JS, Kumar A, Rai AN, Singh DP. Cyanobacteria: A Precious Bio-resource in Agriculture, Ecosystem, and Environmental Sustainability. *Front Microbiol.* 2016;7.
20. Flores E, Herrero A. Nitrogen assimilation and nitrogen control in cyanobacteria. *Biochem Soc T.* 2005;33(1):164-7.
21. Muro-Pastor MI, Reyes JC, Florencio FJ. Ammonium assimilation in cyanobacteria. *Photosynth Res.* 2005;83(2):135-50.
22. Esteves-Ferreira AA, Inaba M, Fort A, Araújo WL, Sulpice R. Nitrogen metabolism in cyanobacteria: metabolic and molecular control, growth consequences and biotechnological applications. *Crit Rev Microbiol.* 2018;44(5):541-60.
23. Flores E, Frías JE, Rubio LM, Herrero A. Photosynthetic nitrate assimilation in cyanobacteria. *Photosynth Res.* 2005;83(2):117-33.
24. Herrero A, Muro-Pastor Alicia M, Flores E. Nitrogen Control in Cyanobacteria. *J Bacteriol.* 2001;183(2):411-25.
25. Sakamoto T, Bryant D. Requirement of Nickel as an Essential Micronutrient for the Utilization of Urea in the Marine Cyanobacterium *Synechococcus* sp. PCC 7002. *Microbes Environ.* 2001;16:177-84.
26. Veaudor T, Cassier-Chauvat C, Chauvat F. Genomics of Urea Transport and Catabolism in Cyanobacteria: Biotechnological Implications. *Front Microbiol.* 2019;10:2052.
27. Valladares A, Montesinos ML, Herrero A, Flores E. An ABC-type, high-affinity urea permease identified in cyanobacteria. *Mol Microbiol.* 2002;43(3):703-15.

28. Herrero A, Flores E. Genetic responses to carbon and nitrogen availability in *Anabaena*. Environ Microbiol. 2019;21(1):1-17.
29. Quintero María J, Muro-Pastor Alicia M, Herrero A, Flores E. Arginine Catabolism in the Cyanobacterium *Synechocystis* sp. Strain PCC 6803 Involves the Urea Cycle and Arginase Pathway. J Bacteriol. 2000;182(4):1008-15.
30. Esteves-Ferreira AA, Cavalcanti JHF, Vaz M, Alvarenga LV, Nunes-Nesi A, Araújo WL. Cyanobacterial nitrogenases: phylogenetic diversity, regulation and functional predictions. Genet Mol Biol. 2017;40(1 suppl 1):261-75.
31. Herrero A, Muro-Pastor AM, Valladares A, Flores E. Cellular differentiation and the NtcA transcription factor in filamentous cyanobacteria. FEMS Microbiol Rev. 2004;28(4):469-87.
32. López-Gomollón S, Hernández JA, Wolk CP, Peleato ML, Fillat MF. Expression of *furA* is modulated by NtcA and strongly enhanced in heterocysts of *Anabaena* sp. PCC 7120. Microbiology. 2007;153(1):42-50.
33. Olmedo-Verd E, Flores E, Herrero A, Muro-Pastor Alicia M. HetR-Dependent and -Independent Expression of Heterocyst-Related Genes in an *Anabaena* Strain Overproducing the NtcA Transcription Factor. J Bacteriol. 2005;187(6):1985-91.
34. Zhao M-X, Jiang Y-L, He Y-X, Chen Y-F, Teng Y-B, Chen Y, *et al.* Structural basis for the allosteric control of the global transcription factor NtcA by the nitrogen starvation signal 2-oxoglutarate. Proc Natl A Sci. 2010;107(28):12487-92.
35. Llácer JL, Espinosa J, Castells MA, Contreras A, Forchhammer K, Rubio V. Structural basis for the regulation of NtcA-dependent transcription by proteins PipX and PII. Proc Natl Acad Sci. 2010;107(35):15397-402.
36. Videau P, Rivers OS, Hurd K, Ushijima B, Oshiro RT, Ende RJ, *et al.* The heterocyst regulatory protein HetP and its homologs modulate heterocyst commitment in *Anabaena* sp. strain PCC 7120. Proc Natl Acad Sci. 2016;113(45):E6984-E92.
37. Zeng X, Zhang C-C. The Making of a Heterocyst in Cyanobacteria. Annu Rev Microbiol. 2022;76(1):597-618.
38. Zhang C-C, Laurent S, Sakr S, Peng L, Bédou S. Heterocyst differentiation and pattern formation in cyanobacteria: a chorus of signals. Mol Microbiol. 2006;59(2):367-75.
39. Khudyakov I, Wolk CP. Evidence that the *hanA* gene coding for HU protein is essential for heterocyst differentiation in, and cyanophage A-4(L) sensitivity of, *Anabaena* sp. strain PCC 7120. J Bacteriol. 1996;178(12):3572-7.
40. Golden JW, Yoon HS. Heterocyst development in *Anabaena*. Curr Opin Microbiol. 2003;6(6):557-63.
41. Muro-Pastor Alicia M, Flores E, Herrero A. NtcA-Regulated Heterocyst Differentiation Genes *hetC* and *devB* from *Anabaena* sp. Strain PCC 7120 Exhibit a Similar Tandem Promoter Arrangement. J Bacteriol. 2009;191(18):5765-74.
42. Flores E, Picossi S, Valladares A, Herrero A. Transcriptional regulation of development in heterocyst-forming cyanobacteria. Biochim Biophys Acta Gene Regul Mech. 2019;1862(7):673-84.
43. Xu X, Risoul V, Byrne D, Champ S, Douzi B, Latifi A. HetL, HetR and PatS form a reaction-diffusion system to control pattern formation in the cyanobacterium *Nostoc* PCC 7120. eLife. 2020;9:e59190.
44. Xu X, Rachedi R, Foglino M, Talla E, Latifi A. Interaction network among factors involved in heterocyst-patterning in cyanobacteria. Mol Genet Genomics. 2022;297(4):999-1015.
45. Videau P, Rivers OS, Tom SK, Oshiro RT, Ushijima B, Swenson VA, *et al.* The *hetZ* gene indirectly regulates heterocyst development at the level of pattern formation in *Anabaena* sp. strain PCC 7120. Mol Microbiol. 2018;109(1):91-104.
46. Huang X, Dong Y, Zhao J. HetR homodimer is a DNA-binding protein required for heterocyst differentiation, and the DNA-binding activity is inhibited by PatS. Proc Natl Acad Sci. 2004;101(14):4848-53.
47. Shvarev D, Maldener I. Roles of DevBCA-like ABC transporters in the physiology of *Anabaena* sp. PCC 7120. Int J Med Microbiol. 2019;309(5):325-30.
48. Kurio Y, Koike Y, Kanesaki Y, Watanabe S, Ehira S. The CRP-family transcriptional regulator DevH regulates expression of heterocyst-specific genes at the later stage of differentiation in the cyanobacterium *Anabaena* sp. strain PCC 7120. Mol Microbiol. 2020;114(4):553-62.
49. Wolk CP, Ernst A, Elhai J. Heterocyst Metabolism and Development. En: Bryant DA, editor. The Molecular Biology of Cyanobacteria. Dordrecht: Springer Netherlands; 1994. p. 769-823.

50. Fresenborg LS, Graf J, Schätzle H, Schleiff E. Chapter 7 - Iron homeostasis of cyanobacteria: advancements in siderophores and metal transporters. En: Singh PK, Kumar A, Singh VK, Shrivastava AK, editores. *Advances in Cyanobacterial Biology*. Academic Press; 2020. p. 85-117.
51. Huertas MJ, López-Maury L, Giner-Lamia J, Sánchez-Riego AM, Florencio FJ. Metals in cyanobacteria: analysis of the copper, nickel, cobalt and arsenic homeostasis mechanisms. *Life (Basel)*. 2014;4(4):865-86.
52. Sevilla E, Sarasa-Buisan C, González A, Cases R, Kufryk G, Peleato ML, *et al.* Regulation by FurC in *Anabaena* Links the Oxidative Stress Response to Photosynthetic Metabolism. *Plant Cell Physiol*. 2019;60(8):1778-89.
53. Peleato ML, Teresa Bes M, Fillat MF. Iron Homeostasis and Environmental Responses in Cyanobacteria: Regulatory Networks Involving Fur. En: de Bruijn FJ, editor. *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria*. 2016. p. 1065-78.
54. Fillat MF. The FUR (ferric uptake regulator) superfamily: Diversity and versatility of key transcriptional regulators. *Arch Biochem Biophys*. 2014;546:41-52.
55. González A, Angarica VE, Sancho J, Fillat MF. The FurA regulon in *Anabaena* sp. PCC 7120: *in silico* prediction and experimental validation of novel target genes. *Nucleic Acids Res*. 2014;42(8):4833-46.
56. González A, Bes MT, Peleato ML, Fillat MF. Expanding the Role of FurA as Essential Global Regulator in Cyanobacteria. *Plos One*. 2016;11(3):e0151384.
57. González A, Valladares A, Peleato ML, Fillat MF. FurA influences heterocyst differentiation in *Anabaena* sp. PCC 7120. *FEBS Lett*. 2013;587(16):2682-90.
58. Sein-Echaluce VC, González A, Napolitano M, Luque I, Barja F, Peleato ML, *et al.* Zur (FurB) is a key factor in the control of the oxidative stress response in *Anabaena* sp. PCC 7120. *Environ Microbiol*. 2015;17(6):2006-17.
59. López-Gomollón S, Sevilla E, Bes MT, Peleato ML, Fillat María F. New insights into the role of Fur proteins: FurB (Al12473) from *Anabaena* protects DNA and increases cell survival under oxidative stress. *Biochem J*. 2009;418(1):201-7.
60. Olivan-Muro I, Sarasa-Buisan C, Guio J, Arenas-Busto J, Sevilla E, Fillat MF. Unbalancing Zur (FurB)-mediated homeostasis in *Anabaena* sp. PCC7120: Consequences on metal trafficking, heterocyst development and biofilm formation. *Environ Microbiol*. 2023.
61. Videau P, Cozy Loralyn M, Young Jasmine E, Ushijima B, Oshiro Reid T, Rivers Orion S, *et al.* The trpE Gene Negatively Regulates Differentiation of Heterocysts at the Level of Induction in *Anabaena* sp. Strain PCC 7120. *J Bacteriol*. 2014;197(2):362-70.
62. Picossi S, Valladares A, Flores E, Herrero A. Nitrogen-regulated Genes for the Metabolism of Cyanophycin, a Bacterial Nitrogen Reserve Polymer: expression and mutational analysis of two cyanophycin synthetase and cyanophycinase gene clusters in heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120. *J Biol Chem*. 2004;279(12):11582-92.
63. Yingping F, Lemeille S, Talla E, Janicki A, Denis Y, Zhang C-C, *et al.* Unravelling the cross-talk between iron starvation and oxidative stress responses highlights the key role of PerR (ahr0957) in peroxide signalling in the cyanobacterium *Nostoc* PCC 7120. *Env Microbiol Rep*. 2014;6(5):468-75.
64. Sarasa-Buisan C, Guio J, Broset E, Peleato ML, Fillat MF, Sevilla E. FurC (PerR) from *Anabaena* sp. PCC7120: a versatile transcriptional regulator engaged in the regulatory network of heterocyst development and nitrogen fixation. *Environ Microbiol*. 2022;24(2):566-82.
65. Stock AM, Robinson VL, Goudreau PN. Two-Component Signal Transduction. *Annu Rev Biochem*. 2000;69(1):183-215.
66. Higo A, Nishiyama E, Nakamura K, Hihara Y, Ehira S. cyAbrB Transcriptional Regulators as Safety Devices To Inhibit Heterocyst Differentiation in *Anabaena* sp. Strain PCC 7120. *J Bacteriol*. 2019;201(17):10.1128/jb.00244-19.
67. Flaherty BL, Van Nieuwerburgh F, Head SR, Golden JW. Directional RNA deep sequencing sheds new light on the transcriptional response of *Anabaena* sp. strain PCC 7120 to combined-nitrogen deprivation. *BMC Genomics*. 2011;12(1):332.
68. Lu H, Wang L, Li S, Pan C, Cheng K, Luo Y, *et al.* Structure and DNA damage-dependent derepression mechanism for the XRE family member DG-DdrO. *Nucleic Acids Res*. 2019;47(18):9925-33.
69. Wang C, Ye F, Kumar V, Gao Y-G, Zhang L-H. BswR controls bacterial motility and biofilm formation in *Pseudomonas aeruginosa* through modulation of the small RNA *rsmZ*. *Nucleic Acids Res*. 2014;42(7):4563-76.



70. Brown NL, Stoyanov JV, Kidd SP, Hobman JL. The MerR family of transcriptional regulators. *FEMS Microbiol Rev.* 2003;27(2-3):145-63.
71. Fang C, Zhang Y. Bacterial MerR family transcription regulators: activation by distortion. *Acta Biochim Biophys Sin.* 2022;54(1):25-36.
72. Fang C, Li L, Zhao Y, Wu X, Philips SJ, You L, *et al.* The bacterial multidrug resistance regulator BmrR distorts promoter DNA to activate transcription. *Nat Commun.* 2020;11(1):6284.
73. Blasco B, Stenta M, Alonso-Sarduy L, Dietler G, Peraro MD, Cole ST, *et al.* Atypical DNA recognition mechanism used by the EspR virulence regulator of *Mycobacterium tuberculosis*. *Mol Microbiol.* 2011;82(1):251-64.
74. Lata S, Mahatha AC, Mal S, Gupta UD, Kundu M, Basu J. Unraveling novel roles of the *Mycobacterium tuberculosis* transcription factor Rv0081 in regulation of the nucleoid-associated proteins Lsr2 and EspR, cholesterol utilization, and subversion of lysosomal trafficking in macrophages. *Mol Microbiol.* 2022;117(5):1104-20.
75. Blasco B, Chen JM, Hartkoorn R, Sala C, Uplekar S, Rougemont J, *et al.* Virulence Regulator EspR of *Mycobacterium tuberculosis* Is a Nucleoid-Associated Protein. *PLOS Pathog.* 2012;8(3):e1002621.
76. Yang Y, Liu C, Zhou W, Shi W, Chen M, Zhang B, *et al.* Structural visualization of transcription activated by a multidrug-sensing MerR family regulator. *Nat Commun.* 2021;12(1):2702.
77. Hayashi K, Kensuke T, Kobayashi K, Ogasawara N, Ogura M. *Bacillus subtilis* RghR (YvaN) represses *rapG* and *rapH*, which encode inhibitors of expression of the *srfA* operon. *Mol Microbiol.* 2006;59(6):1714-29.
78. Pellicer S, Bes MT, González A, Neira JL, Peleato ML, Fillat MF. High-recovery one-step purification of the DNA-binding protein Fur by mild guanidinium chloride treatment. *Process Biochem.* 2010;45(2):292-6.
79. Lu J-j, Shi L, Chen W-l, Wang L. The regulation of HanA during heterocyst development in cyanobacterium *Anabaena* sp. PCC 7120. *World J Microb Biot.* 2014;30(10):2673-80.
80. Fiedler G, Arnold M, Hannus S, Maldener I. The DevBCA exporter is essential for envelope formation in heterocysts of the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol.* 1998;27(6):1193-202.
81. Muro-Pastor AM, Valladares A, Flores E, Herrero A. Mutual dependence of the expression of the cell differentiation regulatory protein HetR and the global nitrogen regulator NtcA during heterocyst development. *Mol Microbiol.* 2002;44(5):1377-85.
82. Frías JE, Flores E. Induction of the Nitrate Assimilation *nirA* Operon and Protein-Protein Interactions in the Maturation of Nitrate and Nitrite Reductases in the Cyanobacterium *Anabaena* sp. Strain PCC 7120. *J Bacteriol.* 2015;197(14):2442-52.
83. Ehira S, Ohmori M. NrrA, a nitrogen-responsive response regulator facilitates heterocyst development in the cyanobacterium *Anabaena* sp. strain PCC 7120. *Mol Microbiol.* 2006;59(6):1692-703.
84. Risser Douglas D, Callahan Sean M. HetF and PatA Control Levels of HetR in *Anabaena* sp. Strain PCC 7120. *J Bacteriol.* 2008;190(23):7645-54.
85. Young-Robbins Shirley S, Risser Douglas D, Moran Jennifer R, Haselkorn R, Callahan Sean M. Transcriptional Regulation of the Heterocyst Patterning Gene *patA* from *Anabaena* sp. Strain PCC 7120. *J Bacteriol.* 2010;192(18):4732-40.
86. Wolk CP, Zhu J, Kong R. Genetic Analysis of Heterocyst Formation. En: Peschek GA, Löffelhardt W, Schmetterer G, editores. *The Phototrophic Prokaryotes*. Boston, MA: Springer US; 1999. p. 509-15.
87. Lechno-Yossef S, Fan Q, Ehira S, Sato N, Wolk CP. Mutations in Four Regulatory Genes Have Interrelated Effects on Heterocyst Maturation in *Anabaena* sp. Strain PCC 7120. *J Bacteriol.* 2006;188(21):7387-95.



# MATERIAL SUPLEMENTARIO



## 9. Material Suplementario

### 9.1. Programa utilizado en las reacciones de PCR

1. Desnaturalización inicial del DNA: 5 min a 95 °C
  2. Desnaturalización del DNA: 30 s a 95 °C
  3. Hibridación del DNA: 30 s a una temperatura de 2°C inferior a la T<sub>m</sub> de la pareja de cebadores utilizados.
  4. Etapa de elongación: 45 s a 72 °C
  5. Elongación final: 10 min a 72°C
  6. Conservación a 4 °C
- } 30 ciclos

Las parejas de cebadores empleadas se recogen en los apartados 9.2 y 9.3 del Material Suplementario.

### 9.2. Cebadores empleados para la amplificación de los genes *all0345* y *alr1976* para el clonaje en el plásmido pET28a

Las secuencias de oligonucleótidos empleadas para amplificar los genes *all0345* y *alr1976* se detallan en la siguiente tabla:

Cebadores	Secuencia (5'→3')	T <sub>m</sub> (°C)
all0345_pET28_up(NdeI)	GGAATTCCATATGtactgtagacagttggatagattaag	60
all0345_pET28_dw(SalI)	TTTGTTCGACtcaagcttgagctactggaaactg	
alr1976_pET28_up(NdeI)	GGAATTCCATATGggaagttatctagatagaaacgc	60
alr1976_pET28_dw(SalI)	TTTGTTCGACttatgggtcttgctcaattcgtc	

**Tabla S-1.** Cebadores empleados para amplificar los genes de interés mediante PCR.

### 9.3. Cebadores empleados para la amplificación de promotores mediante PCR

Las secuencias de oligonucleótidos empleadas para la amplificación de promotores por PCR fueron las siguientes:

Cebadores	Secuencia 5'→3'	Tm (°C)
P0345_up	GATATAACTTGGCAAACGCCTG	63
P0345_dn	TGAAGACAAACGACTTTCTCAC	
P1976_up	ACGCGGATATGGTACAGCAG	60
P1976_dn	ACTAGGCTCGCCAAGCGTTC	
PhenR_up	TTCCTTGAACCAAGATTTGC	61
PhenR_dn	AACAAGCACCCAGTTACCTC	
PhglE_up	TAAGGACAATAGAAGATATCCTAC	59
PhglE_dn	CAATAGGTGTTTGTGTAAAGC	
PhepA_up	CAGTATCTGTAGGGTGCCTATC	44
PhepA_dn	CTTGCTTGCTGAAATAAGTATAG	
PhepB_up	CCCTAGAGGATGAAAATTGAAG	61
PhepB_dn	GGCATTCTCTCATGATATAGAG	
PhepC_up	GCCAATTAAAGTTGTCATGATTAAG	38
PhepC_dn	CCTAACCTTTGTAATATCCGCC	
PhepK_up	CGCTGGTTATGAAGCTGG	48
PhepK_dn	CCCTCAAATCTCCTTACCCTC	
PhetR_up	GGGAAAGTCCTTGTAGGTTAC	52
PhetR_dn	CGCTTGATCAGATCGATG	
PhetC_up	CATAGTTTAATTTCTGTTTGGTG	58
PhetC_dn	CAGATGTCGATGTTGCTCAC	
PntcA_up	GCAATCTATACCCCTATACCC	60
PntcA_dn	CTTATCTTGTGTCACGATCATC	
PnrrA_up	GCCAAAAATTA ACTCTGTG	57
PnrrA_dn	CAACGATTTC AATACAAACC	
Pnir-nrt-nar_up	GGAAGGTGTTGACATCTATATG	60
Pnir-nrt-nar_dn	GAAATTTTCTTCTGGAAACGTG	
PglnA_up	AAGCCTGTTACTGCATCGCGCATTC	56
PglnA_dn	CGAAACAAAGTTGATGAC	
PnifH_up	GGCGTGTGGCTCTTGTTTC	64
PnifH_dn	GGTTGGTTAGTTGGTCGG	
PdevH_up	GGTTTTTATGAGGGGCTAAG	61
PdevH_dn	GATGGAGATTGCATAGGTAAC	
PdevBCA_up	CACCTCGGACACAATCAAAC	63
PdevBCA_dn	GAGATTCCTGAATCGGTCTG	
Pasr1734_up	GTTGATGTGAAAAGCTTTAAAAAC	59
Pasr1734_dn	CATAGTTTTAGAGATTCCTGTGC	

PpatA_up	GGTTTGAGCCAATAACTATACGTG	62
PpatA_dn	GTAATCGGAAGTGTTTTCATGG	
PpatS_up	GCATATTTACGTAAAAAGTTCTATC	58
PpatS_dn	CCGGATAGTTAAATGCAAAAAG	
PhanA_up	GGTGATGACAAGGCTTGG	60
PhanA_dn	CATTTGGAACCTCCTTAATGTTTAC	
PhetN_up	GTTATGAGTGAGTTATCCTGG	60
PhetN_dn	CATTGTAACCTGCTAGTCTCC	
PhetZ_up	GCGTTTAGTTTATCCGCAA	62
PhetZ_dn	CTCAAGCATTGTTGTAGCCG	
PhetC_up	GCTACTAGAAATGAGGAGAGGG	64
PhetC_dn	GTGAGCAACATCGACATCTG	
Ifpkn22_up	AAAGATGAATTACACTGGCG	47
Ifpkn22_dn	CTGCAAACTGTGGCAGAATA	

**Tabla S-2.** Cebadores empleados para amplificar los promotores de los genes de interés mediante PCR.

#### 9.4. Composición de los geles de SDS-PAGE

Reactivo	Gel separador (17 %)	Gel separador (15 %)	Gel concentrador (5 %)
Agua destilada	0,36 mL	0,71 mL	1,4 mL
Tris-HCl 1,5 M, pH 8,8	2,24 mL	2,24 mL	-
Tris-HCl 0,5 M, pH 6,8	-	-	0,25 mL
Acrilamida-Bisacrilamida (30:0,8 p/v)	3,4 mL	3 mL	0,33 mL
SDS 10 % (p/v)	30 µL	30 µL	10 µL
PSA 10 % (p/v)	20 µL	20 µL	10 µL
TEMED	10 µL	10 µL	10 µL

**Tabla S-3.** Composición de los geles de poliacrilamida con SDS (SDS-PAGE) utilizados en la electroforesis en condiciones desnaturalizantes.

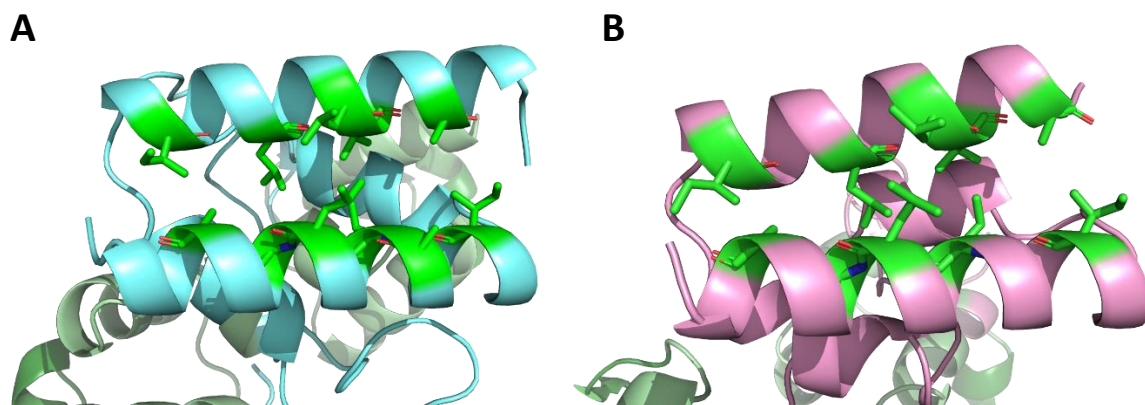
#### 9.5. Composición de los geles empleados en los ensayos de retardo en gel

Reactivo	Volumen
Agua destilada	4,79 mL
Acrilamida:Bisacrilamida (30:0,8 p/v)	2 mL
Glicerol 50 % (v/v)	1,4 mL
Running buffer 10x (30,28 g/L Tris-HCl, 142 g/L glicina pH 8,5)	0,93 mL
PSA 10 % (p/v)	50 µL
TEMED	30 µL

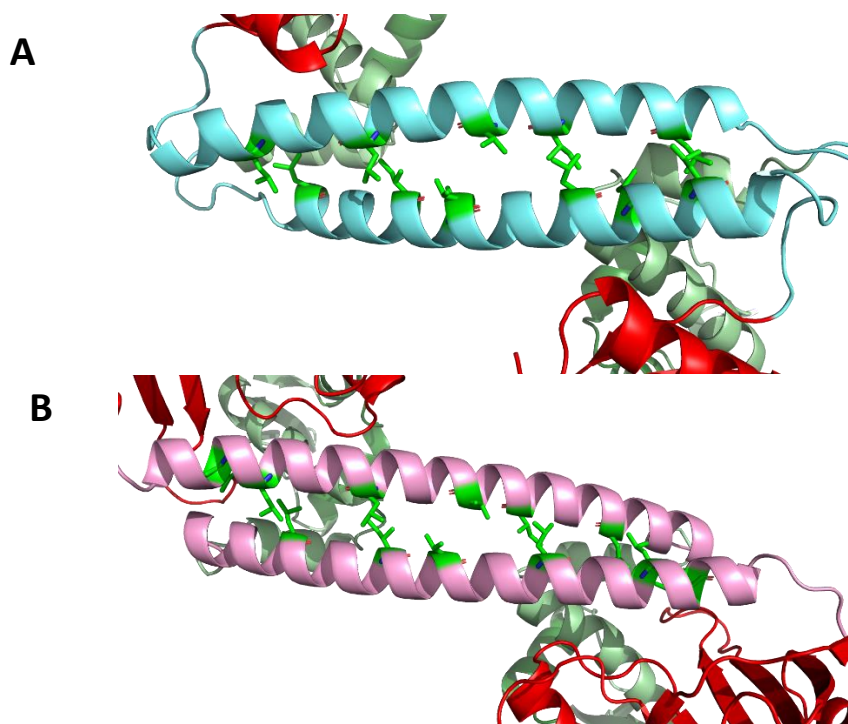
**Tabla S-4.** Composición de los geles de poliacrilamida al 6 % utilizados en los ensayos de retardo en gel



### 9.6. Visualización de los residuos hidrofóbicos presentes en las hélices implicadas en la dimerización de Alr1976 y All0345



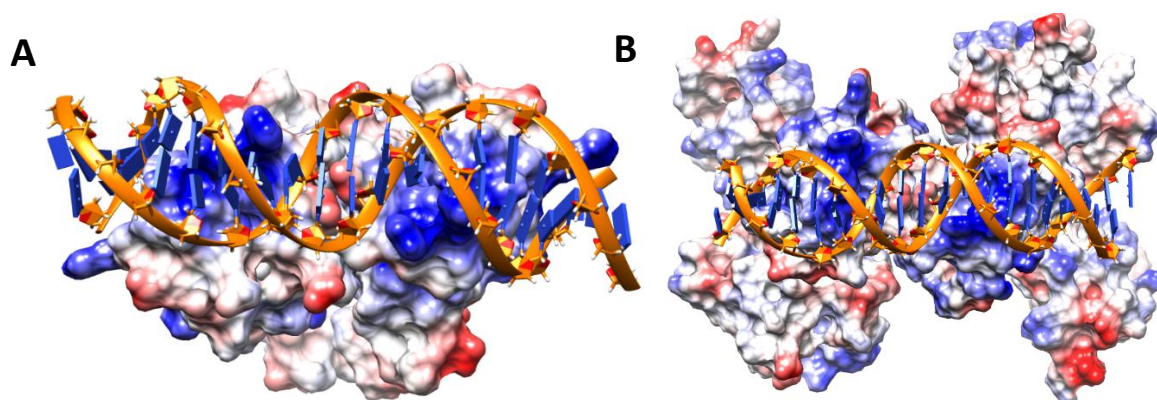
**Figura S-1.** Visualización de los residuos hidrofóbicos presentes en las hélices  $\alpha$  implicadas en la dimerización de Alr1976 modelado en base a BswR (PDB: 4O8D) (A) y en base a EspR (PDB:3QF3) (B).



**Figura S-2.** Visualización de los residuos hidrofóbicos presentes en las hélices  $\alpha$  implicadas en la dimerización de All 345 modelado en base a EcmR (PDB: 6XLK) (A) y en base a BmrR (PDB:1R8E) (B).

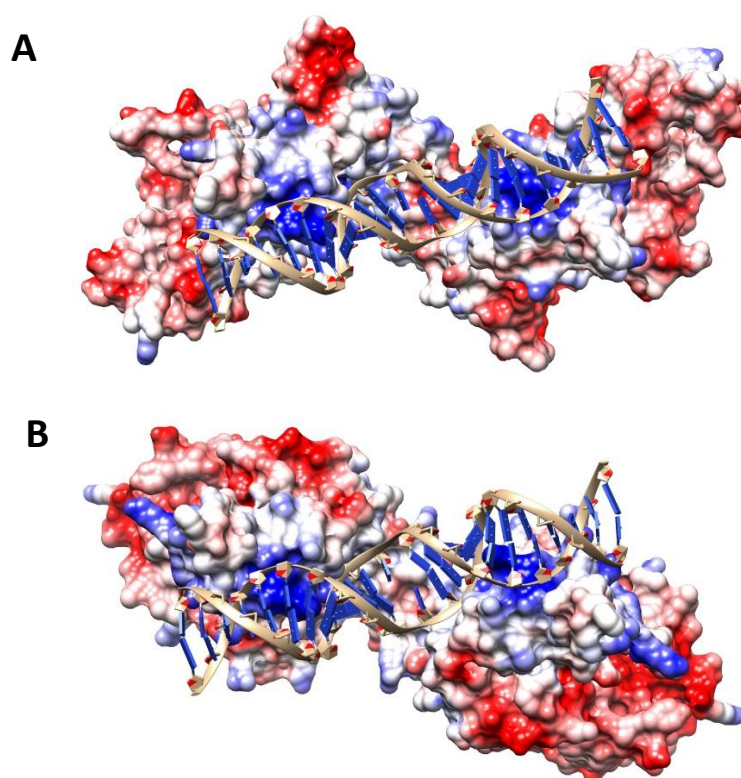
### 9.7. Modelo del complejo DNA-regulador mostrando la carga electrostática

En la figura S-3, se representa el modelo generado del complejo DNA-Alr1976 mostrando la superficie coloreada según el potencial electrostático.



**Figura S-3.** Modelo del complejo DNA-Alr1976 modelado en base a BswR (PDB: 4O8B) (A) y a EspR (PDB: 3QF3) (B) mostrando la carga. En azul está representada la carga positiva, mientras que, en rojo, se muestra la carga negativa.

En la figura S-4, se representa el modelo generado del complejo DNA-All0345 mostrando la superficie coloreada según el potencial electrostático



**Figura S-4.** Modelo del complejo DNA-All0345 modelado en base a EcmR (PDB: 6XLK) (A) y a BmrR (PDB: 1R8E) (B) mostrando la carga. En azul está representada la carga positiva, mientras que, en rojo, se muestra la carga negativa.

## 9.8. Secuencias empleadas para realizar los alineamientos de las proteínas Alr1976 y AlI0345 con otras secuencias de proteínas de cianobacterias

A continuación, se muestran las secuencias de las proteínas de cianobacterias empleadas para realizar el alineamiento múltiple y el árbol filogenético de la proteína Alr1976 de *Anabaena*:

```
>Anabaena_sp._PCC7120
MGSYLDIERLASLVRKKRGSRLRETSVEIGNVSPSTLSRVESGRMPDMETFLALCNWLQVPPAELFRTTEEDQP
DTPEAIAIQLRADKKLDPAIANALASLVKAAYRDLSSQDDELKQDP

>Aetokthonos_hydrillicola
MVRARRGRPLHEIALEIGNISPSTLSRIENGKMPDMETFLLLCDWLQVPVTELIKSTESFSVNNSDAITMQIQA
DKNLDPANALAVLIKAAYKQDVAVFHAYESSNKDIRRHSTN

>Calothrix_membranacea_FACHB-236
MGSYLDIERLASLVRKQGRSRLRETAADIGKVSPSTLSRVESGKIPDMETFLALCNWLQVPPAELFRSTDADQL
NTPEAIAIQLRADKNLDPAIANALASLVKAAYQDLSQKNSELKQDP

>Chamaesiphon_polymorphus
MANYLDMEKLANLVNRKRLTRGLRETAKEIGNVSPSTISRVENGKTPDMETFLALCDWLAVPAAELIKNTEDEKA
PNTPEAITIQLRADRNLDPAIANALASLVTAAYKDLSQKNNEVASDRNQDEE

>Chlorogloea_sp._CCALA_695
MGSYLDIERLASLVRSKRGSCGLRETALIIGNVSPSTISRVENGKIPDMDTFLALCDWLEVPSELIKNTEDEHL
DTPEAISIQLRADKNLDPAIANALASLVKAAYKDLSQKNKEE

>Chondrocystis_sp._NIES-4102
MKNYLDTERLANLVNRKGRNRLRETAKEIGNVSPSTISRVENAKTPDMDTFLALCDWLEIPPAELIKNAEEEEEE
TLNTPEAITIQLRADKNLDPAIALALASLVKAAYKDLSKNKDE

>Komarekiella_sp.
MESYLDIERLASLVRKKRGSRLRETATDIGNVSPSTLSRVESGKMPDMETFVALCNWLQVPPAELFRTSKEEQQL
DTPEAIAIQLRADKNLDPAIANALASLVKAAYRDLSSQDNGELKQDP

>Merismopedia_sp._SIO2A8
MSYHLDLNRLATLRSRRGSLGLRAAAKEIGNVSTSTLSRVENGKMPDMETFLALCDWLEVPASLIKNDENKSE
ASRLDSSAS

>Mojavia_pulchra_JT2-VF2
MGYNLDTHLLAAMMKTKRGRKGLRDTAKEIGNISSSTLSRVENGKMPDMETFLLLCDWLQVSPAELFKKTEESEEA
SSGSNTLETIEIQLRASKKLDPAVANALAEVLKAAYRDNSQQSDTQQPD

>Scytonema_hofmannii
MGSYLDMNQLATIVRTRGTRGLRETAKEIGNISPSTLSRVENGKMPDMETFLLLCDWLEVPAAELIKNTEEAGN
PSVLETPDAIAIQLRADKNLDPAIANALASLVKAAYHDLSQSNTNPKQES

>Symploca_sp._SIO2B6
MSYHLDLNRLATLRSRRGSLGLRAAAKEIGNVSTSTLSRVENGKMPDMETFLALCDWLEVPASLIKNDENKSE
ADESNLDASEALAIQLRADKNLDPATANALAEVLKAAYANLPKRIDALHTDSSQSDAAE

>Tolypothrix_sp._PCC_7910
MGSYLDIEKLASLVRKKRSSRLRETAEEIGNVSPSTLSRVESGKMPDMETFLALCNWLQVPPAELFKTTDADQL
NTPEAIAIQLRADKNLDPAIANALASLVKAAYRDLSSQDDELKQGS

>Trichormus_sp._M33_DOE_039
MEYHLDTNLLAARVKTGRGKGLRDTAREIGNISSSTLSRVENGKMPDMETFLLLCNWLQVSPAELFKKTEESEV
PDSNTLEAIEIQLRASKKLDPAVANALAEVLKAAYRDNAQQSDE
```

A continuación, se muestran las secuencias de las proteínas de cianobacterias empleadas para realizar el alineamiento múltiple y el árbol filogenético de la proteína All0345 de *Anabaena*:

>Anabaena\_sp.PCC\_7120

MYCRQLDRLSNPYHLQRSKMLKIGDFSKLSQVSVKALRLYDQMGLLKPIKVDNFTSYRYSSAEQLPRLNRILAF  
KDLGFSLEQIAKLLDENLPPEQIRGMLRLKQGEIQRLVEAEQARLIRVEARLKQIEQEDSMPNYEVVKKVAPIQ  
VASIRQILPDNPSIGQLYGEISEYLAQNGVKAGDYAGIWHDPGYKDTDIDAEAVISIEGSIKNERIKIYELPG  
SETTACLIHHGSYETLAQAYATLVSWIEANGYNITAPNREVYIIGGNEQNDSYVTELQFPVAQA

>Brasilonema\_sp.\_CT11

MLKIGDFSKLSHVTVKALRLYDQLGLLKPSYVDHFTGYRYYLADQLIRLNRLILAFKDLGFSLEQIAKLLNENLPP  
AEIRGMLRLKQVELQQLVEEQQAQLLRVEARLKQIEQEDTMPNYEVVLKTVEPIKVASIREVLDPDPSIGRLYPE  
LLEYFKQYEVKMGDYWGAIWHDPVYKESEIDGEAVISFEGEILSTASVKVYELPSIEKMATVIHHGSYNTINQAY  
AVLPAWIETNGYKIIGSNREVYIIRGNEQDNESYVTEVQFPVTKA

>Calothrix\_parasitica

MLKIGDFSKLSQVTVKALRLYDQLGLLKPSRIDEFTGYRYYSVQQLPRLNRILAFKDLGFSLEQIAKLLDENLPP  
TEIRGMLRLKQVELEQLLVEEQARLFRVEARLKQIEREDKMPNYEVVKKIEPITVAGIRDIMPNYSQMSQLYEE  
LFGYFQQHGIKVDNYCAGILHDPYKESDVDVEAVVSIDGTVPINERIKVYELPGYEQAACVVHQSNTINQAY  
TALLSWIESNGYKITDSNREVYIVGGNEQDNDSYVTELQFPINLK

>Chlorogloeopsis\_fritschii

MLKIGDFSKLSQVTVKALRLYDQLGLLKPIYVDEFTSYRYYCAEQQLPRLNRILALKDLGFSLEQIAKLLDENVPS  
AEIRGMLRLKQAEQLRLVEEQARLLRVEARLKQIEEESIPKYEIVLKKVEPIKVASIREVLPNFAAVPRLYDQL  
FDTLQQQGIQEPSYCAGIWHDSYKESDWDWEAVASAEFVSSNKRKVKYELLGFEEKMACAVHHGSYDTSNQAYA  
ALISWIEANGYKIIGSNREVYIIGGNEQNDSYVTEIQFPVARA

>Drouetiella\_hepatica\_Uher\_2000/2452

MLRIGDFSRLAQVSVRTLRYEELSLLKPIEVDRFTDYRYYSVDQLPRLNRVLALKDLGFSLNQIKDLLDSDLTL  
LELRAMLTVRQAEIEQQQLQEGMTKLTRVEARLRQIEREGKLADYEVVLKRVEPKVVAATRQIVPTVTAILPYRGK  
MYAEVYTWNQHRIKAGEPEMAIYHNAEYTEKDIDMEMAVIDPAAVPSALYPLKEQVYFYELPLVQTMASLTRY  
GKLSDIGQGLIELFTWMGENGYSADGAYREIHLFGREFEQVQDQWECNPQLDFDAVVVELQIPVKRLS

>Fischerella\_thermalis

MLKIGDFSKLSQVSVKALRLYDQLGLIKPVDVDNFTGYRYSSAEQLPRLNRILALKDLGFSLEQIAKLLDENLPS  
TEIRGMLRLKQAEQLHLVEEQARLLRVEARLKQIEQEDNMPNYEVVKKVEPIKVASIREKLDPDPSVGQLYDE  
LLTYFNQQGVKKGTYCAGIWHDPGYKDSNVDGEAVISIEKDINTEIRIRIYELPRFEKMACLVHHSYNTLNQAY  
AALVSWIEGNGYQIIGPNREVYITGGNEQDNECYITEVQFPVTKA

>Gloeocapsa\_sp

MLKIGDFSKLSHVSIALRLYDQMGLLKPVHVDEFTGYRYAAHQQLPRLNRILALKDLGFSLEQIAKLLNEISPA  
QIQGMLQLKHAELQRLVAEEQARLQVAARLQQMEEDNMPDYDVVVKTVTPIKVAAIREIIPSFHAVAGLYDEL  
EYLRQCQGVKESYSCAGIWHDSAYKDTDIDWEVAIAVEEVYGSDRKIKVYELPRVEMACVVHNGSYNTISRAYAAL  
PAWIEASGYKITGNSREVYIVGGNQDDESYVTEVQFPVKKLD

>Gloeocapsopsis\_dulcis

MLKIGDFSKLSHVSIALRLYDQMGLLKPVHIDRFTGYRYSAHQQLPRLNRILALKDLGLSLEQIAKLLDEISPA  
QIQGMLRLKQAEQLQVLVEEQARLQVAARLQQIEQENKMPEYDVVLKTVAPIKVAAREIIPSFQVVAQLCDEL  
LEYLQRQGVKGSSYFAGIWHDTAYKNTDIDWEVAISVEEVCSSDRKIKVYELPCIEMACVVHNGSYNTINRAYAAL  
PVWIEANGYKIAGSNREVYIVGGDRQDDESYVTEVQFPLELRNC

>Goleter\_apudmare

MLKIGDFSKLSQVTVKALRLYDQLGLLKPVHVDNFTGYRYYSADQLPRLNRILVLKDLGFSLEQIAKLLDEKLPP  
EQIRGMLRLKQAEQLQLIAEQSRLVRVEVQLRQIEQEDSMPNYDIVLKKLEPIKVASIREILPNFATVAKLYDE  
LSNYIQQQGVKEFNYSAGIWHDPYKESDIDWEFVISINNFIPSNERVQIYELPGYETTACAVHQSYNTIKQAY  
NALASWIEANGYKISGASREVYIVGGVEQDNDSYVTEVQFPVPSQA

>Hydrococcus\_sp

MLKISDFAQLSRISSKTLRLYDRMGLLKPTAVDSFTGYRYYSTSQLPRLNRILVLKELGFSLEQIVRLLDENLSV  
DEIRGMLRLKQYEQRLLEEDRIRLLRIEMRLLEIEQEGKMPYEIVLKSVEPQLVATTGLVIPNYQDCGEIFDR  
FFDRAYEYISQHGKIKVCGIAIYHDTLRDRDIPVEAAAPIYEKIASDERVWVYELPAIETIACAVHRGSFATL  
GQAYNALLGWVEKNGYQIVGSTREVYLEYERGGDQSQYVTEVQVPVENN

>Iphinoe\_sp

MLKIGDFSLSHVTVKALRLYDQLGLLKPSHVDHFTGYRYYLADQLPRLNRILAFKDLGFSLEQIAKLLNENLSS  
AEIRGMLRLKQAEQLQRLVEEEQTRLLRVEARLKQIEQEDTMPNYEVVLKTVEPIKVASIREILPNYASIGRLYSE  
LLEYFKQHVKVKEGDYWGAIWHDVPYKESDVDGEAVISFEGEIPSTERLKVYELPVVEKMATVVHNGSYSSINQAY  
AALPAWIETNGYKIIGSNREVIYVGGHEQDNESYVTEVQFPPIAQA

>Kovacikia\_minuta

MFKIGDFSRLSRVSIRTLRLYDEMDDLKPVCDVQFTGYRYYSSASQLAQLNRIVALKDLGFSLEQIAKLLDEKIPP  
TQIRGMLRLKQAEQLQHVELEQARLARVEARLSQIEYEDFVSKYDIVLKKSESQTVAAMRDILPTCDRIAQLHSE  
LENYLSAQSSITIVGFPQTIWHD AEYRRQD VDAEAVVPINQSLTTTERFRTYQLPEIQMACVIHHGSYDSIFQAF  
NALSQWIEANHFLIVGPNREVIYVLPQALEAKVSQVVRTVNNRSDSVVEVQFPVPIPEHAQA

>Leptolyngbya\_sp

MFKIGDFSRLSRVSRALRLYDQMGLLKPVHVDVFTGYRYYSAEQLSRLNQIVAFKDLGFSLEQIIELLNDKIPP  
AQIRGMLRLKQAEVQQLIEGEQARLKRIETRLQIEQSDFTSLYNVTLKQVDAQSVVSIRDILPICSNIKHLHIE  
LEDALQHQNMGIVGFPQTLWHD AEYRVNNVD AEAIAPVKQSFSGSERVKYFVLPVEVQQMACTIHHGSYDTLVQAF  
NALLYWIEANNYQVAGPNREIYLPKPKITNIDFLGHADCTVIEIQFPICDL

>Mastigocoleus\_testarum

MLRISDFAQLSRVSPKALRLYDRMGLLKPIKVDSFTGYRYYSAIQLPRLNRILVFKELGFSLEQIAKLLDENIPA  
EEIRGMLRIKHLEIQORLLEDQVRLARVEMRLSEIEQEKKMPNYEVILKPVESQLVAVTVGVIPNYQDCGP IFER  
LFGEAYAYVSRHNLKQSGNGIS IYHDTKLDRYIPVEAAVPIYEKIP SNEKVLVYELPGVEDMACTVHQGSFASL  
GQAYNILLEWVEKNGYRIIGSTREIYLQYEQGGDESQYVTEVQIPVEKI

>Merismopedia\_glauca

MFRIGEFSQLGQVSVRMLRHYDNLGLLKPAKVDRFTDYRYYTIEQLPRLQRILALKDLGLSLEQIVSLLDNELSP  
EQLRSMMLLKQAEQAQLEDTRSRNLNRVVARLRQIEQEEQPYGYDVTLSKSVNYAEIASIRQLVPTIEDMGT YRHQ  
ASNLLYEWLEAQKLTPKGAEIVLYHLAEYSEVDIDMEFGIILDESIEVRS LKQFLPQSPVKIRELTSV SQMASLV  
HSGMLQDIAQGITALYQWIGTNNLIPVGA FREIHLFGKETDRVLDEPVVVELQIPVTSKRSHL

>Microcoleus\_sp

MLKIGDFSLSQVSIKALRLYDRMGLLKPTVVDYDTAYRYYSSASQLPRLNRILAFKDLGFSLEQIPLPLDDKISP  
VQIRGMLRLKQAEQLQHQHIGETARLARVEARLRQIEQEDDMPDYEVILKTIQIPITVASIRETL PKYAAVGS L FNE  
LYDYLHRNEVKLG DYCLA IWHDAGYKESD VDAEAI VSVDKLLPEGDR I K VYELPGCDRAACAIHHGSYNTITQSY  
NYLLTWIENNGYRIVGANREVIYQGGQE QDNESYVTEIQFPVEKV

>Pleurocapsa\_sp

MFKIGDFSLSQSIKALRLYDRMGLLKPFHIDPYTNYRYYSAKQLPRLNRILAFKDLGFSLEQIRKLL EDDIPL  
EQIRGMLRLKQAEQLKQIESEQLKLN RVKARIKQIEEENTMRNYDVVVKVEPIKVASIREILPNYSAVGELFNE  
LIGYLQYQKITKFDYCAAIWHDQ EYKESD VDAEAVLVIDTALAETDR I K VYELPRYDEIACI IHQGSYLT LNQSY  
ENLLAWIETNGYQIIAPNRELYIQGGQE QDNESYITEIQFPVQKV

>Pseudanabaena\_galeata

MLKIGDFSLSQVTIKALRLYDEMGLLKPI SVDRFTAYRFYSATQLPRLNRILAFKDLGFSLEQISQLLNEEISP  
EQMRGMLRLKQADLQQQIEQE QGR LVRVAARLKQIEQEDVMSDYEVV I KKFEPITVASIREVL PN YAAIAKL FQE  
LYQYLAQQGVTKFEYDAGI WHD SGYRES DIDGEAVVSVPTSVKGTDR I K VYELQGYEAIACV VHHGSYSTLNNGY  
QHLLTWIENNGYRCIGANREFYIQGGAELDNESYITEIQFP IEKI

>Rivularia\_sp

MLKIGDFSLSQVTVKALRLYDQLGLLKPADIDEFTGYRYYSVQQLPRLNRILAFKDLGFSLEQIAKLLDDNLPP  
VEIRGMLRLKQAEHLKLEEEQARLLRVETRLKQIEREDKMPNYEVV I K KIEPIKVAGIRDIVPNY GAMSQLYEE  
LFGYFQQHG I K VNDY CAGILHDPEYKESD VDEAIITFDGAVPINKRIK VYELPGYKQAACLIYQGSYSKINQAY  
TALLSWIENNGYKIIDS NREVIYVIGGNEQDNDSYVTELQFPISVK

>Scytonema\_sp

MLKIGDFSLSHVTVKALRLYDQLGLLKPSHVDHFTGYRYY SADQLPRLNRILAFKDLGFSLEQIAKLL EENLSA  
TEIRGMLRLKQAEQLQRLVEEEQARLLRVEARLKQIEEDTMPNYEVVLKTVEPIKVASIREILPDYPSVGS L YDEL  
LK YFKQHA AKEGDYCGAIWHD PGYKESD VDAEAVLSFEGDIPSTERVKVYELPGIEKMATVIHHGSYKTINQAYA  
ALPTWIETNGYKIIGPNREVIYIGGNEQDNESYVTQVQLPVAKA

>Synechococcus\_sp

MLKIGDFSRLGQVSVRMLRHYDQMGLIKPAYVDPDSEYRYYSLQLPRLHRILALKDLGLSLEQIETALQTD LAP  
GQLKEMLALKQAE LSQQVAETQIRLQRVASRLAQLETEAVVSPYEVVLRAVPERVLASIRRVPTLMDMPAYRCG  
ASDELYAWLDKQHLQPCGP EMVLYHFVEYTETD VMEFGIPVQPRPGLTFRDPIQLASLP AVDQMACVIHTGSLW  
DVGLAITALFTWMGRNGYASAGSFREVHLFGKETEK TAREPVVVELQLPVVSQSIHLQRASQ SREA

>Tolypothrix\_sp

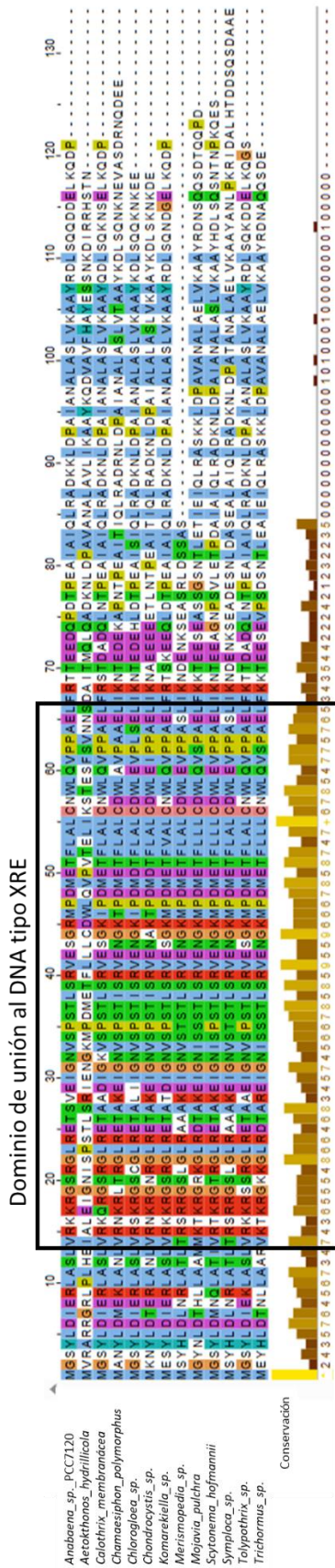
MLKIGDFSKLSTVTKALRLYDQLGLLKPAHIDHFTGYRYYSANQLPRLNRILALKDLGFSLEQIAKLLEDNLPS  
DQIRGMLRLKQAEQLRLVEEEHARLLRVEVRLKQIEKDDTMPKYEVVLKKVDPIKVVSIREILPDYPSVGRLYDE  
LLGYLNQQGAKEGSYCAAIWHDPGYKESDVGAEVFSIESSIQGNERIKVYELPGWETIACLVHNGSYNTINQAY  
AALVSWIEANGYKIIGPNREVYIIGGNQQDNDTYVTEVQFPVSKA

>Trichormus\_sp

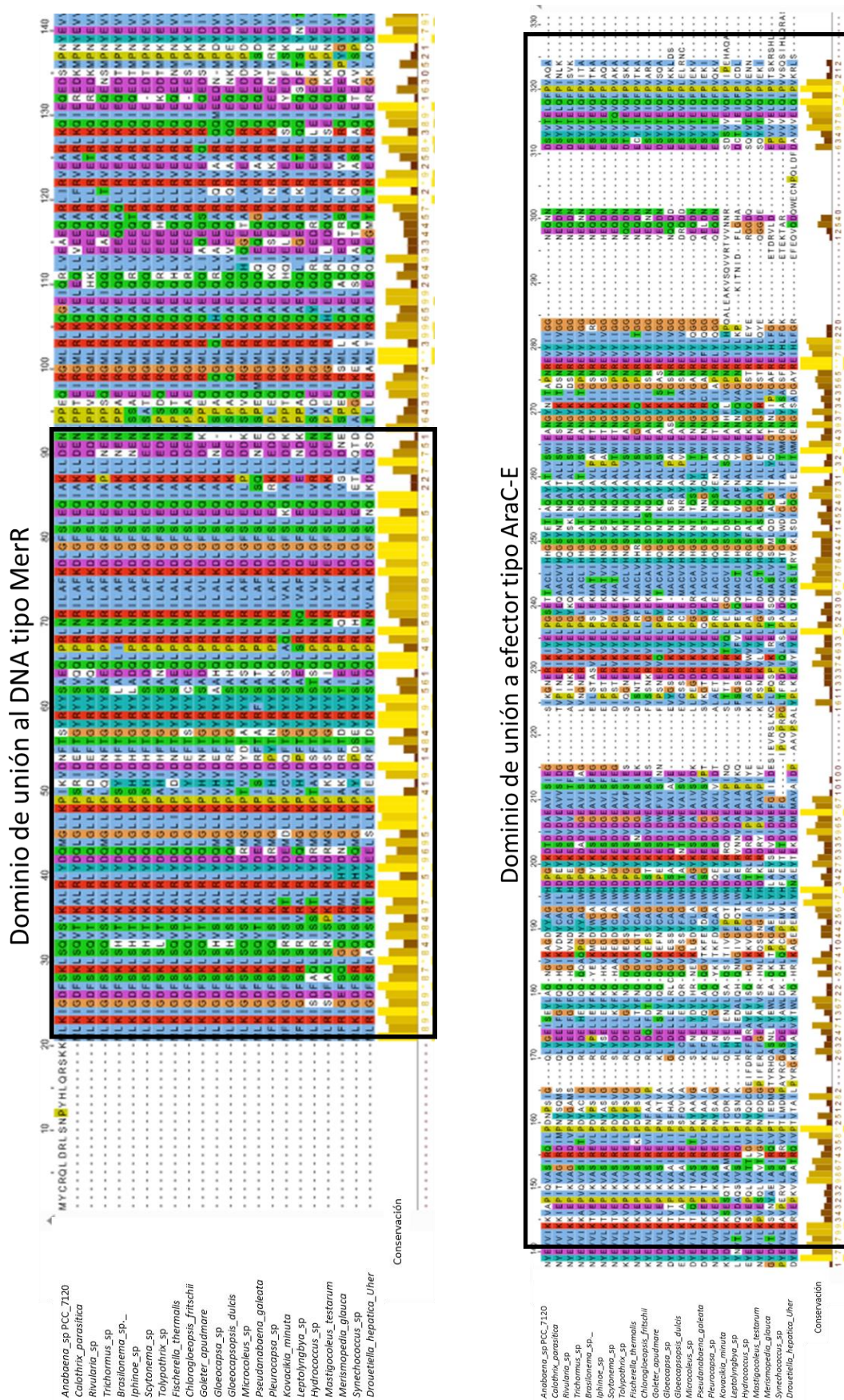
MLKIGDFSKLSTVSVKALRLYDQMGLLKPLQVDNFTGYRYYSAEQLPRLNRILAFKDLGFSLEQIAPLLNENLPP  
SQIRGMLRLKQAEIQQLVEAEQARLTRVEARLKQIEQENSMPNYEVVIKKAEPVQVVSIRETLPDYACIGRLYDE  
LHEYLQQQQAQPGNYYAGIWHDPGYKDADVDGEAVISIAGLVNGNERIKVYELPGLEAIACLIHHGSYATLSQAY  
ATLVSWIEANGYKITGPNREVYIVGGNEQNNEYSVTEIQFPVITA



### 9.9. Alineamiento de las secuencias de las proteínas Alr1976 y All0345 con secuencias homólogas de cianobacterias



**Figura S-5.** Alineamiento múltiple de Alr1976 con secuencias homólogas en otras cianobacterias. La primera secuencia corresponde a Alr1976 de *Anabaena*. También se muestra la conservación de cada residuo Visualizado en Jalview.





## 9.10. Secuencias empleadas para realizar los alineamientos de las proteínas Alr1976 y Alr0345 con otras secuencias de proteínas de otros procariotas diferentes de cianobacterias

A continuación, se muestran las secuencias de las proteínas de otros procariotas empleadas para realizar el alineamiento múltiple y el árbol filogenético de la proteína Alr1976 de *Anabaena*:

```
>Alr1976 [Anabaena sp. PCC 7120]
MGSYLDIERLASLVRKKRSGRLRETSVEIGNVSPSTLSRVESGRMPDMETFLALCNWLQVPPAELFRRTTEEDQPT
PEAIAIQLRADKKLDPAIANALASLVKAAAYRDLSSQQDDELKQDP
>SinR [Bacillus subtilis]
MIGQRIKQYRKEKGYSLELAEKAGVAKSYLSSIERNLQTNPSIQFLEKVS AVLDSVHTLLDEKHETEDYDGLDS
EWEKLVRDAMTSGVSKKQFREFLDYQKWRKSQKEE
>SutR [Escherichia coli]
MENLARFLSTTLKQLRQQRGWSLSRLAEATGVSKAMLGQIERNESSPTVATLWKIATGLNVPFSTFISPPQSATPS
VYDPQQQAMVITSLFPYDPQLCFEHFSIQMASGAISESTPHEKGVIEHVVIDGQLDLCDVDGEWQTLNCGEGVRFA
ADVTHIYRNGGEQTVHFSLSLIHYPRS
>BcrR [Enterococcus faecalis]
MEFNELQLRLTGKNTLQEQLAELVSRRTAISKEWESGKGYPNMESLKCISKFFSVTIDELLSGEELITLAETENR
SNLKKIYNYIYGILDMMAVAFIFLPLYGNSVGGYVYAVNLLSFTATTPFNLA VYWSAFAALIIIGIGKIISTHLDK
EKWGGIATKCSLTITALAVCFFAAAREPYITVLVFLLLIGKIFVWIKQMGMK
>ClgR [Mycobacterium smegmatis]
MTALLREVIGDVLNRNARTDQGRTLREVSDAARVSLGYLSEVERGRKEASSELLSAICDALDVPLSRVLT DAGESMA
RREHDAREAEQVAVANLGRIDAATKVVIPQVSMAVA
>DdrOC [Deinococcus deserti]
MKLHERLRELRLSERGLRLKDVAEVAQISVPYLSDLERGRTNPSLETQLTAGAYNITVHDLLEGVEFYGASTEGAL
PKGLSDLIADPTLGPQITPDWVRTLSRIELRGKRPRDKQDWYEIYLHLKRILS
>RghR [Bacillus subtilis]
MESFGEQLRALREERKLTVNQLATYSGVSAAGISRIENGKRGVPKPATIKKLAEALKIPYEGLMYKAGYIEEVHEA
RAPYETKCKLLEKAEAYDLKNLALLENEKWQYLNKEDLLMLDHYFSFISDEAKKRSADD
>DicA [Escherichia coli]
METKNLTIGERIRYRRKNLKHQTSRLAKALKISHVSVSQWERG DSEPTGKNLFALSKVLQCSPTWILFGDEDKQPT
PPVEKPVALSPKELELLELFNALPESEQDTQLAEMRARVKNFNKLFEE LLKARQRTNKR
>PuuR [Escherichia coli]
MSDEGLAPGKRLSEIRQQGLSQRRAEELSGLTHSAISTIEQDKVSPAISTLQKLLKVYGLSLSEFFSEPEKPDEP
QVVINQDDLIEMGSQGVSMKLVHNGNPNRTLAMI FETYQPGTTTGERIKHQGEEIGTVLEGEIVLTINGQDYHLVA
GQSYAINTGIPHSFNTSAGICRIISAHTPTTF
>SlrR [Bacillus subtilis]
MIGRIIRLYRKRKGYSINQLAVESGVSKSYLSKIERGVHTNPSVQFLKKVSATLEVELTELFD AETMMYEKISGGE
EEWRVHLVQAVQAGMEKEELFTFTNRLKKEQPETASYRNRKLTESNIEEWKALMAEAREIGLSVHEVKSFLKTKGR
>Xre [Bacillus subtilis]
MIGGRKSLRGKRTQEEIASHIGVSRARYSHYENGRSEPDYDTLQKLADYFQVTTDYLLTGDKKSDDDMFSDPDL
QLAYRDMQDFSPESKQQAIEFINYLKEKEKNRKPKNK
>ImmR [Bacillus subtilis]
MSLGKRLKEARQKAGYTQKEAAEKL NIGNNNLSNYERDYRDPD TDTLLKLSNLYNVSTDYLLGKDEVSKKNETDLL
NKTINEAIQELKDEDTLLFMNDGEFDEETARLVKKALKNGIKFIDELKKKE
>AnsR [Bacillus subtilis]
MNLDRLTELRKKKNWSLQYTADLLGIAKSTYAGYESGYRRPSLEALAMLADLFDTTCD ELLGREKQKQTAPQAIEL
ATWSSLDFTISVDGQPLSEDEIIQLITFIRTKRKVQEELS
>BswR [Pseudomonas Aeruginosa]
MLGTRLKAARIRAGYSQKQLGMLVGMDEF SASARMNQYERERHSPNMRTSEQLAMVLQVPMAYLYCPEDELAELIL
KVSSLTPEFKKELTRFIEQLLAAQGSTRQPVTRSEL
```

>EspR [Mycobacterium tuberculosis]

MSTTFAARLNRLFDTVYPPGRGPHTSAEVIAALKAEGITMSAPYLSQLRSGNRTNPSGATMAALANFFRIKAAAYFT  
DDEYYEKLDKELQWLCTMRDDGVRRIAQRAHGLPSAAQQKVLDRIDELRRAEGIDA

A continuación, se muestran las secuencias de las proteínas de otros procariotas empleadas para realizar el alineamiento múltiple y el árbol filogenético de la proteína Alr1976 de *Anabaena*:

>All0345 [Anabaena sp. PCC7120]

MYCRQLDRLSNPYHLQRSKKMLKIGDFSKLSQVSVKALRLYDQMGLLKPIKVDNFTSYRYYSAEQLPRLNRILAFK  
DLGFSLEQIAKLLDENLPPEQIRGMLRLKQGEIQRLVEAEQARLIRVEARLKQIEQEDSMPNYEVVKKVAPIQVA  
SIRQILPDNPSIGQLYGEISEYLAQNGV  
KAGDYYAGIWHDPGYKDTDIDAEAVISIEGSIKNERIKIYELPGSETTACLIHHGSYETLAQAYATLVSWIEANG  
YNITAPNREVYIIGGNEQNND SYVTELQFPVAQA

>CueR [Escherichia coli]

MNISDVAKITGLTSKAIRFYEEKGLVTPPMRSENGYRTYTQQHLNELTLLRQARQVGFNLEESGELVNLFNDPQRH  
SADVKKRTLEKVAEIERHIEELQSMRDQLLALANACPGDSDADCP I IENLSGCCCHHRAG

>TnrA [Bacillus subtilis]

MTTEDHSYKDKKVISIGIVSELTGLSVRQIRYEEERKLIYPQRSSRGTRKYSFADVERLMDIANKREDGVQTAEIL  
KDMRKKEQMLKNDPQVRKKMLEGQLNAHFRYKNR

>SoxR [Escherichia coli]

MEKKLPRIKALLTPGEVAKRSGVAVSALHFYESKGLITSIRNSGNQRRYKRDVLRVVAI I KIAQRIGIPLATIGEA  
FGVLPEGHTLSAKEWKQLSSQWREELDRRIHTLVALRDELDGCIGCGCLSRSDCPLRNPGDRLGEEGTGARLLEDE  
QN

>ZntR [Escherichia coli]

MYRIGELAKMAEVPDPTIRYEEKQMMHEVTEGGFRLYTESDLQRLKFIRHARQLGFSLESIRELLSIRIDPEH  
HTCQESKGIVQERLQVEEARIAELQSMQSRSLQRLNDACCGTAHSSVYCSILEALEQGASGVKSGC

>MlrA [Escherichia coli]

MALYTIGEVALLCDINPVTLRWQRRYGLLKQPRTDGGHRLFNADADIDRIEIKRWIDNGVQVSKVKMLLSNENVD  
VQNGWRDQQETLLTYLQSGNLHSLRTWIKERGQDYPAQTLTTHLFIPLRRLQCQQPTLQALLAILDGVLINYYIAI  
CLASARKKQGDALVVGWNIQDTTRLWLGEWIASQQGWRIDVLAHSLNQLRPELFEGRITLLVWCGENRTSAQQQQQL  
TSWQEQGHDIPLGI

>GlnR [Bacillus subtilis]

MSDNIRRSMPFLPIGIVMQLTELSARQIRYEEENGLIFPARSEGNNRRLFSFHDVDKLL EIKHLIEQGVNMAGIKQI  
LAKAAEPEQKQNEKTKKPMKHDLSDDDELRLQLLNELMQAGRFQRGNTFRQGDMSRFFH

>BluR [Escherichia coli]

MAYYSIGDVAERCGINPVTLRWQRRYGLLKQPQRSEGGHRLFDEEDIQRIEEIKRWISNGVPVGKVKALLETTSQD  
TEDDWSRLQEEMMSILRMANPAKLARI IISLGREYPVDQLINHVVLPVRQRLVLDHNTSRIMSSMFDGALIEYAAT  
SLFEMRRKPGKEAILMAWNVEERARLWLEAWRLSLSGWHISVLADPIESPRPELFPTQTLIVWTGMAPTRRQNEL  
QHWGEQGYKVIFHAP

>CarA [Myxococcus xanthus]

MTLRIRTIARMTGIREATLRAWERRYGFPRPLRSEGNRYRVYSREEVEAVRRVARLIQEEGLSVSEAIQAQVKTEPP  
REQPEAERLRERFWSSVGALEGDEVTRVLDDAQTVMDVEAYCDGFLPLREMGVRLDVAREHLASALIRQLRQV  
YDALSPAPAGPRALLACPSGDHHEGGLLVGLIHLKRKGWRVTMLGADTPAAALQGACVQVRPDVVALSFVRARAPE  
EFASVLEDALRACAPFPVVVGGGLGAREHLKAIFSLGAQYAESSEELVAIWNQVRNAQNRP

>AdhR [Bacillus subtilis]

MNIAQVAKQFGLTAATLRYERVGLIPPVKRKDSGIRDYDEEDIKWIEFIKCMRNAGLSIEALIEYTTLFTEGDRT  
VEARKNILADERQRLIEKRKEIDETIKRLDTKIKDYDGKLRENEAKLKS RPKTESLHGSVEQRR

>MerR [Staphylococcus aureus]

MGMKISELAKACDVNKETVRYERYKGLIAGPPRNESGYRIYSEETADRVRFIKRMKELDFSLKEIHLLFGVVDQDG  
ERCKDMYAFTVQKTKEIERKVQGLLR IQRLLLEELKEKCPDEKAMYTCPI IETLMGGPDK

>BmrR [Bacillus subtilis]

MKESYYSIGEVSPLANVSIKALRYDKIDLFKPAYVDPDTSYRYYTDSQLIHLDLIKSLKYIGTPLEEMKKAQDLE  
MEELFAFYTEQERQIREKLDFLSALEQTISLVKKRMKRQMEYPALGEVFLDEEEIR I IQTEAEGIGPENVLNASY

SKLKKFIESADGFTNNSYGATFSFQPYTSIDEMTYRHIFTPVLTNKQISSITPDMEITITIPKGRYACIAYNFSPEH  
YFLNLQKLIKYYIADRQLTVVSDVYELIIPHYSPKKQEEYRVEMKIRIAE

>Mta [*Bacillus subtilis*]  
MKYQVKQVAEISGVSIRTLHHYDNIELLNPSALTDAGYRLYSDADLERLQQILFFKEIGFRLDEIKEMLDHPNFDR  
KAALQSQKEILMKKKQRMDEMIQTIDRTLSSVDGGETMKNRDLFAGLSMKDIEEHQQTYADEVRKLYGKEIAEETE  
KRTSAYSADDWRTIMAEFDSIYRRIAARMKHGPDDAEIQAAVGAFRDHICQYHYDCTLDIFRGLGEVYITDERFTD  
SINQYGEGLAAFLREAIIIYCDHQENPRP

>HmrR [*Sinorhizobium medicae*]  
MNIGEASKVSGVSSKMIRYYEQIGLISPAVRTASSYRTYGDNDVHTLRFIRRRARDLGFSVEQIKELLALWRDRSRA  
SSDVKAVALEHIAELERKIAAIQDMTRTLKHLASHCHGDGRPDCPIIEEMAKGGGAakteINPRFGVASLK

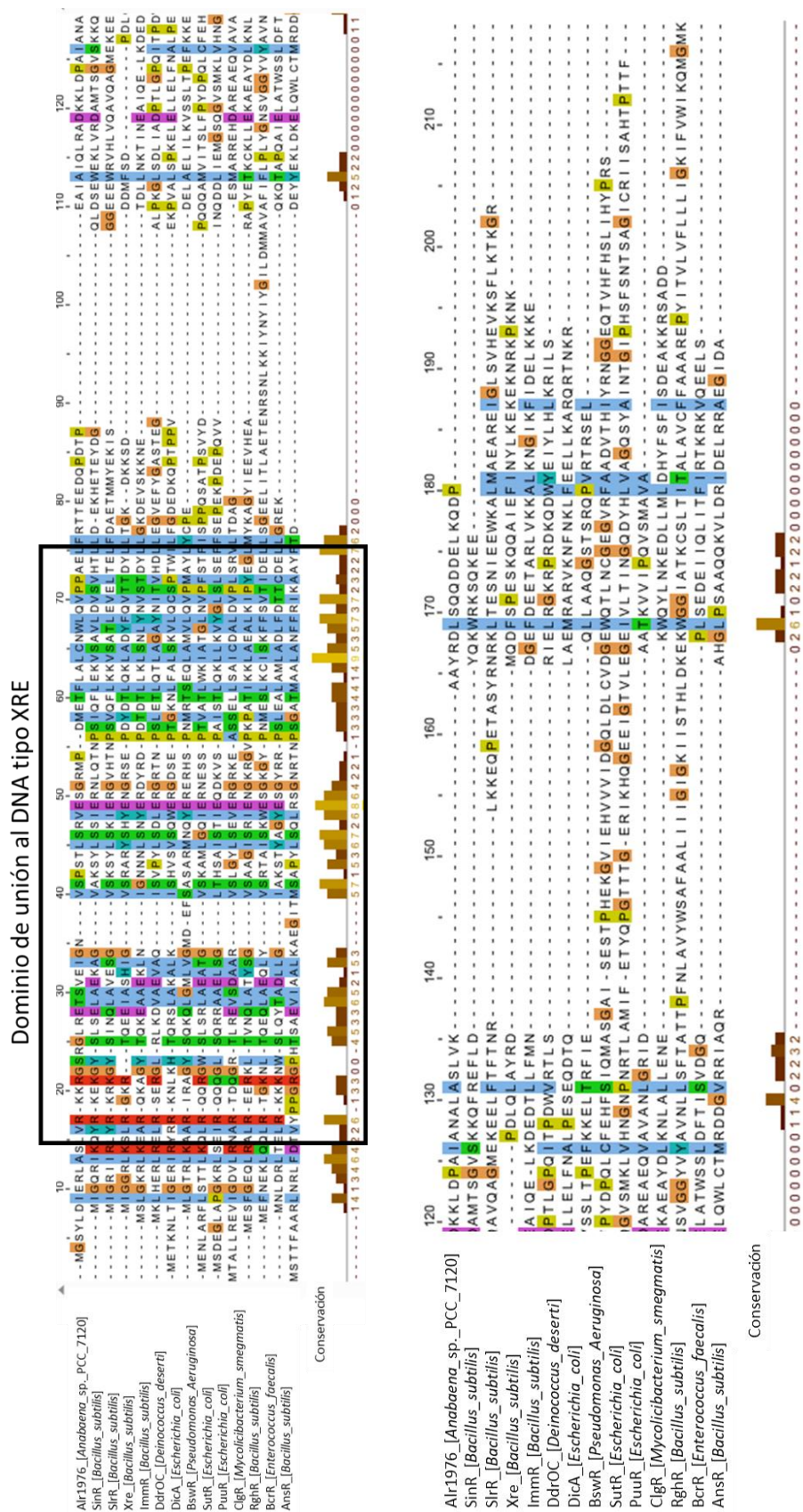
>TipA [*Streptomyces coelicolor*]  
MSYSVGQVAGFAGVTVRTLHHYDDIGLLVPSERSHAGHRRYSADLDRLQQILFYRELGFPLDEVAALLDDPAADP  
RAHLRRQHELLSARIGKLQKMAAAVEQAMEARSMGINLTPEEKFEVFGDFDPDQYEEEVREWRGNTDAYRQSKEKT  
ASYTKEDWQRIQDEADELTRRFVALMDAGEPADSEGAMDAAEDHRQGIARNHYDCGYEMHTCLGEMYVSDERFTRN  
IDAAKPGLAAYMRDAILANAVRHTP

>BltR [*Bacillus subtilis*]  
MSEDVKKYFTTGEFSKLCRVKKQTLFHYDEIGLFSPEIKKENG YRYYSYHQFETFQVISLKFELGVPLKEIKCLIK  
GKTPDKILHVLKEKSIEIDKKINELKQLQTILQTKVTLTEQALETDFSSISFEYLNEETFMLS RKTNLNPERKYVA  
AISELIHEVQQYELDEGYPIGGIFAREQILEKDFYNYSYFYIKVKDGAENIN YHVRPKGLYAVGYEIGGNTEEAYR  
RIIEFIERNGMQIGENAYEEYMLDEM VVDGYENTYAKILLQVKEV

>Nola [*Bradyrhizobium elkanii*]  
MSKATPRRRRWRIGELAEATGVTVRTLHHYEHTGLLAASERTDGGHRMYDRESVQRVHQIRALRELGFSLHEIRKA  
MEGTTSLIDLRLKHLERIELQVARATLLRDLRDMTTGSEIQVSVDLPATLDAMSKVQTRSQTSRCTCKLAIERE  
ERWRRIRDELRCMDRGEHPCGERAKAVAVAAARLLISEIAGADSRVSTILKVLARLSAPRSLAGWDP

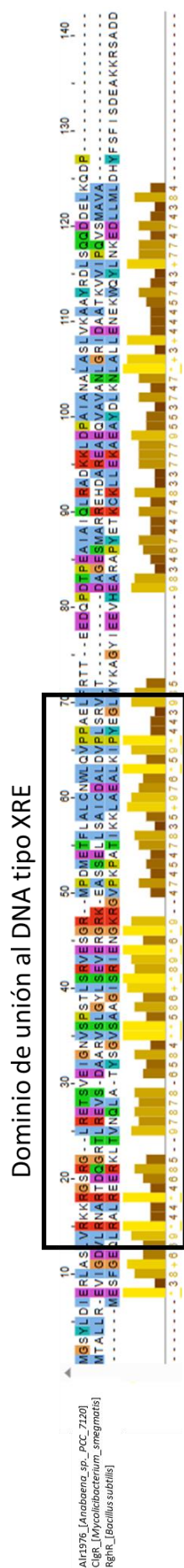
>EcmrR [*Escherichia coli*]  
SQIGLFSKICRVTIKTLHYYNKIGLLVPAYINPDNGYRFYTS DQLMKFHQIASLRQLGFTTITEIVTLTQDENSCHI  
IERRRLEIQKQIRDMADMLSRINHYLQHKKKERIMLYQAALKEIPECIVYSKRFIVPDFSSYIKLIPPIGQEV MKA  
NPGLTTLTPAYCFTLYHDKEYKEKNMDVEFCEAVNDFGKNEGNIIFQVIPAITAVTVIHKGPYDSL RNAYIYLMQW  
VEDNGYLLTNSPRESYIDGIWNKQDSAEWMTEIQFPVEKV

### 9.11. Alineamiento de las secuencias de las proteínas Alr1976 y All0345 con secuencias homólogas de otros procariotas



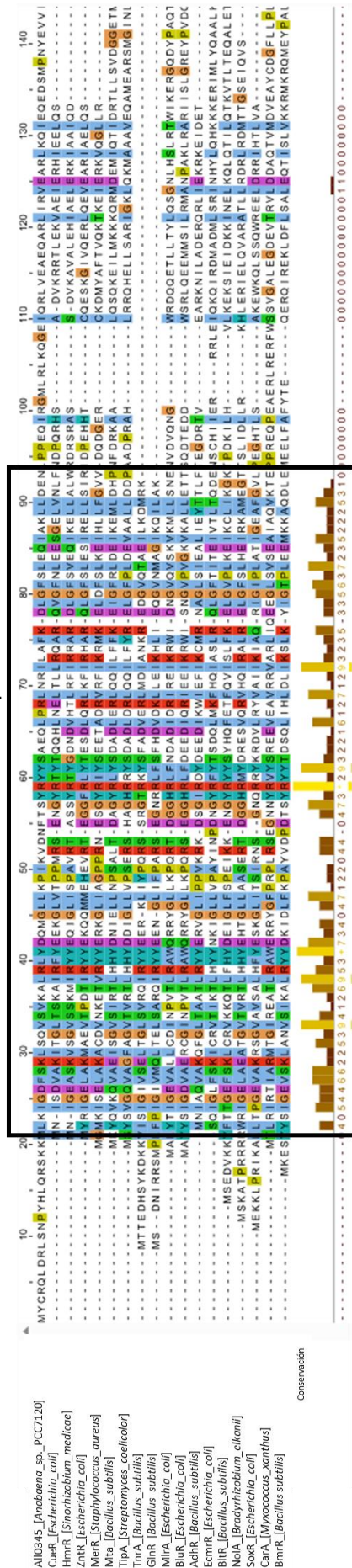
**Figura S-7.** Alineamiento múltiple de Alr1976 con secuencias de reguladores de la familia XRE en otros procariotas diferentes de cianobacterias. La primera secuencia es la de Alr1976 de *Anabaena*. También se muestra la conservación de cada residuo. Visualizado en Jalview.



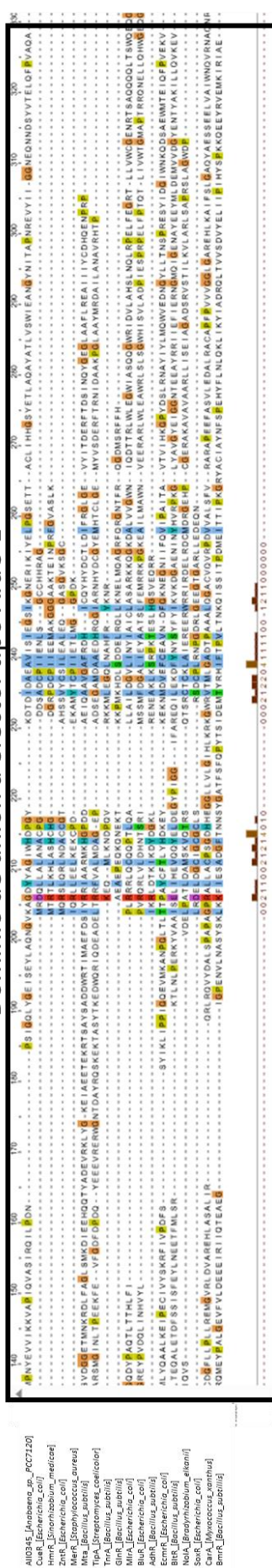


**Figura S-8.** Alineamiento múltiple de Alr1976 con las secuencias filogenéticamente más cercanas de reguladores de la familia XRE en otros procariontes en diferentes cianobacterias. La primera secuencia es la de la proteína Alr1976 de *Anabaena*. También se muestra la conservación de los residuos. Visualizado en Jalview.

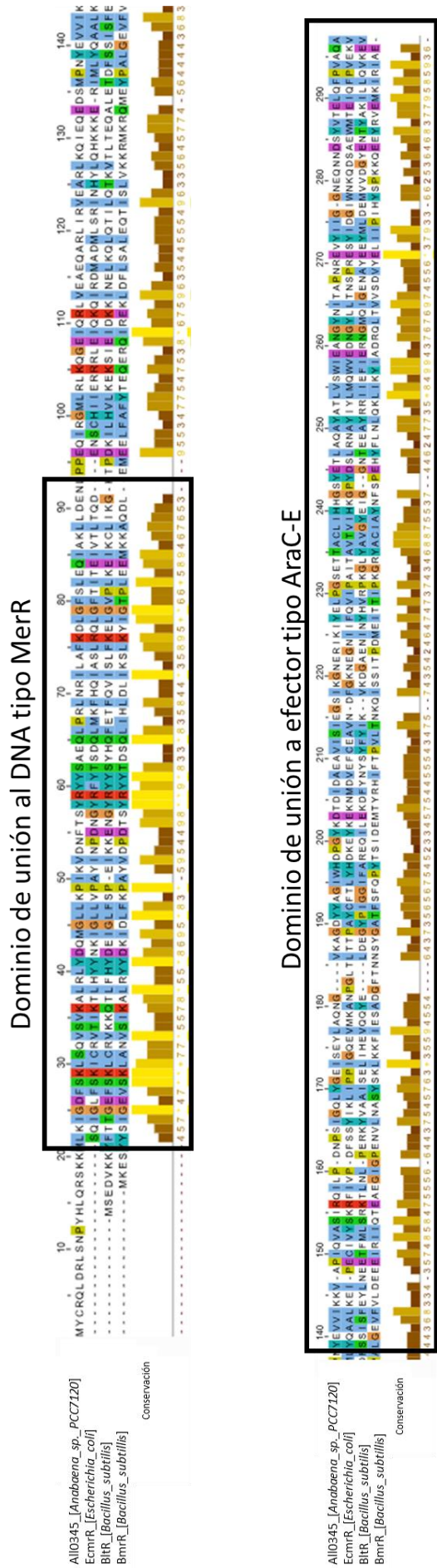
Dominio de unión al DNA tipo MerR



Dominio de unión a efector tipo AraC-E



**Figura S-9.** Alineamiento múltiple de A10345 con secuencias de reguladores de la familia MerR en otros procariontes diferentes de cianobacterias. La primera secuencia corresponde a A10345 de *Anabaena*. También se muestra la conservación de cada residuo. Visualizado en Jalview.



**Figura S-10.** Alineamiento múltiple de AlI0345 con las secuencias filogenéticamente más cercanas de reguladores de la familia MerR en procariotas diferentes de cianobacterias. La primera secuencia corresponde a AlI0345 de *Anabaena*. También se muestra la conservación de cada residuo. Visualizado en Jalview.